

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 13, 2003, 09:31:40 ; Search time 13.6351 Seconds
(without alignments)
65.165 Million cell updates/sec

Title: US-09-290-049A-1
Perfect score: 119
Sequence: 1 ANDHSLLEAMSNDPTPLMD 21

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA: *
1: /cgn2_6/ptodata/2/1aa/5A.COMB.pep.*
2: /cgn2_6/ptodata/2/1aa/5B.COMB.pep.*
3: /cgn2_6/ptodata/2/1aa/6A.COMB.pep.*
4: /cgn2_6/ptodata/2/1aa/6B.COMB.pep.*
5: /cgn2_6/ptodata/2/1aa/PTUS.COMB.pep.*
6: /cgn2_6/ptodata/2/1aa/backfile1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	119	100.0	1475	3 US-09-007-999-2	Sequence 2, Appl1
2	119	100.0	1475	3 US-09-210-361-2	Sequence 2, Appl1
3	119	100.0	1475	4 US-09-740-274-2	Sequence 2, Appl1
4	110	92.4	1375	3 US-09-210-361-4	Sequence 4, Appl1
5	110	92.4	1375	4 US-09-740-274-4	Sequence 4, Appl1
6	75	63.0	545	4 US-09-604-957-4	Sequence 4, Appl1
7	75	63.0	1430	3 US-09-008-172-2	Sequence 2, Appl1
8	75	63.0	1430	3 US-09-210-361-6	Sequence 6, Appl1
9	75	63.0	1430	4 US-09-740-274-6	Sequence 6, Appl1
10	71	59.7	523	4 US-09-604-957-5	Sequence 5, Appl1
11	63	52.9	1577	2 US-08-793-824-2	Sequence 2, Appl1
12	60	50.4	584	4 US-09-604-957-6	Sequence 6, Appl1
13	60	50.4	2057	4 US-09-499-203-2	Sequence 2, Appl1
14	49	41.2	535	4 US-09-604-957-7	Sequence 7, Appl1
15	49	41.2	1278	4 US-09-604-957-3	Sequence 3, Appl1
16	46	38.7	336	4 US-09-198-452A-662	Sequence 662, App
17	46	38.7	596	4 US-09-252-991A-21255	Sequence 21255, A
18	45	37.8	236	4 US-09-154-750A-80	Sequence 80, Appl
19	45	37.8	385	1 US-08-361-920-23	Sequence 23, Appl
20	45	37.8	385	1 US-08-479-939-23	Sequence 23, Appl
21	45	37.8	385	1 US-08-483-432-23	Sequence 23, Appl
22	44.5	37.4	130	4 US-09-328-352-4952	Sequence 4952, Ap
23	44.5	37.4	236	4 US-09-252-991A-31380	Sequence 31380, A
24	44	37.0	234	4 US-09-134-001C-4514	Sequence 4514, Ap
25	44	37.0	384	3 US-09-311-170-2	Sequence 2, Appl1
26	44	37.0	647	4 US-09-134-001C-5458	Sequence 5458, Ap
27	44	37.0	749	4 US-09-562-737-96	Sequence 96, Appl

ALIGNMENTS

28	43	36.1	2465	2 US-08-596-291-3	Sequence 3, Appl1
29	43	36.1	2465	3 US-09-100-804-3	Sequence 1, Appl1
30	43	36.1	2466	3 US-09-080-855-12	Sequence 12, Appl1
31	43	36.1	2466	4 US-09-566-076-12	Sequence 12, Appl1
32	43	36.1	2466	5 PCT-US94-09943-2	Sequence 2, Appl1
33	43	36.1	2465	3 US-09-290-640-46	Sequence 46, Appl1
34	42	35.3	211	4 US-09-107-532A-5029	Sequence 5029, Ap
35	42	35.3	259	1 US-08-277-231A-3	Sequence 3, Appl1
36	42	35.3	259	2 US-08-473-750-6	Sequence 6, Appl1
37	42	35.3	259	2 US-08-477-326-6	Sequence 6, Appl1
38	42	35.3	288	4 US-09-252-991A-23168	Sequence 23168, A
39	42	35.3	989	2 US-08-070-301-16	Sequence 16, Appl1
40	42	35.3	1428	4 US-09-252-991A-30731	Sequence 30731, A
41	41.5	34.9	454	4 US-09-252-991A-28000	Sequence 28000, A
42	41.5	34.9	1048	4 US-09-171-699-10	Sequence 10, Appl1
43	41	34.5	244	4 US-09-252-991A-22373	Sequence 22373, A
44	41	34.5	270	4 US-09-323-872A-38	Sequence 38, Appl1
45	41	34.5	296	4 US-09-328-352-7769	Sequence 7769, Ap

```

RESULT 1
US-09-007-999-2
; Sequence 2, Application US/09007999
; Patent No. 6087559
; APPLICANT: Nichols, Scott E.
; TITLE OF INVENTION: Substitutes for Modified Starch and
; FILE REFERENCE: 0356D
; CURRENT FILING DATE: 1998-01-16
; EARLIER APPLICATION NUMBER: 08/478,704
; EARLIER FILING DATE: 1995-06-07
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 1475
; TYPE: PRT
; ORGANISM: Streptococcus mutans
US-09-007-999-2

Query Match      100.0%; Score 119; DB 3; Length 1475;
Best Local Similarity 100.0%; Pred. No. 1e-10;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 ANDHSLLEAMSNDPTPLMD 21
Db      481 ANDHSLLEAMSNDPTPLMD 501

RESULT 2
US-09-210-361-2
; Sequence 2, Application US/09210361
; Patent No. 6284479
; GENERAL INFORMATION:
; APPLICANT: Nichols, Scott E.
; TITLE OF INVENTION: Substitutes for Modified Starches and
; FILE REFERENCE: 0357CR
; CURRENT FILING DATE: 1998-12-11
; EARLIER APPLICATION NUMBER: 09/007,999
; EARLIER FILING DATE: 1998-01-16
; EARLIER APPLICATION NUMBER: 08/478,704
; EARLIER FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 09/009,620
; EARLIER FILING DATE: 1998-01-20
; EARLIER APPLICATION NUMBER: 08/485,243
; EARLIER FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 09/008,172

```

```

; EARLIER FILING DATE: 1998-01-16
; EARLIER APPLICATION NUMBER: 08/482,711
; EARLIER FILING DATE: 1995-06-07
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 1475
; TYPE: PRT
; ORGANISM: Streptococcus mutans
US-09-210-361-2

Query Match          100.0%; Score 119; DB 3; Length 1475;
Best Local Similarity 100.0%; Pred. No. 1e-10;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 ANDHLSILEAWSNDNTPYLHD 21
Db      481 ANDHLSILEAWSNDNTPYLHD 501

RESULT 3
US-09-740-274-2
; Sequence 2, Application US/09740274
; Patent No. 6465203
; GENERAL INFORMATION:
; APPLICANT: Nichols, Scott E.
; TITLE OF INVENTION: Glucan-containing Compositions and Paper
; FILE REFERENCE: 0357CRD
; CURRENT APPLICATION NUMBER: US/09/740,274
; CURRENT FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: 09/210,361
; PRIOR FILING DATE: 1998-12-11
; PRIOR APPLICATION NUMBER: 09/007,999
; PRIOR FILING DATE: 1998-01-16
; PRIOR APPLICATION NUMBER: 08/478,704
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: 09/009,620
; PRIOR FILING DATE: 1998-01-20
; PRIOR APPLICATION NUMBER: 08/485,243
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: 09/008,172
; PRIOR FILING DATE: 1998-01-16
; PRIOR APPLICATION NUMBER: 08/482,711
; PRIOR FILING DATE: 1995-06-07
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 1475
; TYPE: PRT
; ORGANISM: Streptococcus mutans
US-09-740-274-2

Query Match          100.0%; Score 119; DB 4; Length 1475;
Best Local Similarity 100.0%; Pred. No. 1e-10;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 ANDHLSILEAWSNDNTPYLHD 21
Db      481 ANDHLSILEAWSNDNTPYLHD 501

RESULT 4
US-09-210-361-4
; Sequence 4, Application US/09210361
; Patent No. 6284479
; GENERAL INFORMATION:
; APPLICANT: Nichols, Scott E.
; TITLE OF INVENTION: Substitutes for Modified Starches and
; FILE REFERENCE: 0357CR
; CURRENT APPLICATION NUMBER: US/09/210,361
; CURRENT FILING DATE: 1998-12-11
; EARLIER APPLICATION NUMBER: 09/007,999
```

```

; EARLIER FILING DATE: 1998-01-16
; EARLIER APPLICATION NUMBER: 08/478,704
; EARLIER FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 09/009,620
; EARLIER FILING DATE: 1998-01-20
; EARLIER APPLICATION NUMBER: 08/485,243
; EARLIER FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 09/008,172
; EARLIER FILING DATE: 1998-01-16
; EARLIER APPLICATION NUMBER: 08/482,711
; EARLIER FILING DATE: 1995-06-07
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 1375
; TYPE: PRT
; ORGANISM: streptococcus mutans
US-09-210-361-4

Query Match          92.4%; Score 110; DB 3; Length 1375;
Best Local Similarity 95.2%; Pred. No. 2.8e-09;
Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 ANDHLSILEAWSNDNTPYLHD 21
Db      507 ANDHLSILEAWSNDNTPYLHD 527

RESULT 5
US-09-740-274-4
; Sequence 4, Application US/09740274
; Patent No. 6465203
; GENERAL INFORMATION:
; APPLICANT: Nichols, Scott E.
; TITLE OF INVENTION: Glucan-containing Compositions and Paper
; FILE REFERENCE: 0357CRD
; CURRENT APPLICATION NUMBER: US/09/740,274
; CURRENT FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: 09/210,361
; PRIOR FILING DATE: 1998-12-11
; PRIOR APPLICATION NUMBER: 09/007,999
; PRIOR FILING DATE: 1998-01-16
; PRIOR APPLICATION NUMBER: 08/478,704
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: 09/009,620
; PRIOR FILING DATE: 1998-01-20
; PRIOR APPLICATION NUMBER: 08/485,243
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: 09/008,172
; PRIOR FILING DATE: 1998-01-16
; PRIOR APPLICATION NUMBER: 08/482,711
; PRIOR FILING DATE: 1995-06-07
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 1375
; TYPE: PRT
; ORGANISM: streptococcus mutans
US-09-740-274-4

Query Match          92.4%; Score 110; DB 4; Length 1375;
Best Local Similarity 95.2%; Pred. No. 2.8e-09;
Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 ANDHLSILEAWSNDNTPYLHD 21
Db      507 ANDHLSILEAWSNDNTPYLHD 527

RESULT 6
US-09-604-957-4
; Sequence 4, Application US/09604957
; Patent No. 6486314
```



```
; GENERAL INFORMATION:
; APPLICANT: VAN GEEL-SCHUTTEN, GERRITDINA HENDRIKA
; APPLICANT: DIJKHUIZEN, LUBBERT
; APPLICANT: RAHMOUI, HAKIM
; APPLICANT: LEER, ROBERT-JAN
; TITLE OF INVENTION: NOVEL GLUCOSYLTRANSFERASE AND NOVEL GLUCAN
; FILE REFERENCE: BO 43388
; CURRENT APPLICATION NUMBER: US/09/604,957
; CURRENT FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: 00201871.1
; PRIOR FILING DATE: 2000-05-25
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 545
; TYPE: PRT
; ORGANISM: Streptococcus mutans
; US-09-604-957-4
```

```
Query Match          63.0%; Score 75; DB 4; Length 545;
Best Local Similarity 71.4%; Pred. No. 0.00043;
Matches 15; Conservative 1; Mismatches 5; Indels 0; Gaps 0;
```

```
QY      1 ANDHSLILEAWSNDNTPYLHD 21
Db      75 AINHLILEAWSNDNDPQYKND 95
```

```
RESULT 7
; US-09-008-172-2
; Sequence 2, Application US/09008172
; Patent No. 6127602
; GENERAL INFORMATION:
; APPLICANT: Nichols, Scott E.
; TITLE OF INVENTION: Substitutes for Modified Starch and
; FILE REFERENCE: 0358D
; CURRENT APPLICATION NUMBER: US/09/008,172
; CURRENT FILING DATE: 1998-01-16
; EARLIER APPLICATION NUMBER: 08/482,711
; EARLIER FILING DATE: 1995-06-07
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 1430
; TYPE: PRT
; ORGANISM: Streptococcus mutans
; US-09-008-172-2
```

```
Query Match          63.0%; Score 75; DB 3; Length 1430;
Best Local Similarity 71.4%; Pred. No. 0.0014;
Matches 15; Conservative 1; Mismatches 5; Indels 0; Gaps 0;
```

```
QY      1 ANDHSLILEAWSNDNTPYLHD 21
Db      495 AINHLILEAWSNDNDPQYKND 515
```

```
RESULT 8
; US-09-210-361-6
; Sequence 6, Application US/09210361
; Patent No. 6284479
; GENERAL INFORMATION:
; APPLICANT: Nichols, Scott E.
; TITLE OF INVENTION: Substitutes for Modified Starches and
; FILE REFERENCE: 0357CR
; CURRENT APPLICATION NUMBER: US/09/210,361
; CURRENT FILING DATE: 1998-12-11
; EARLIER APPLICATION NUMBER: 09/007,999
; EARLIER FILING DATE: 1998-01-16
; EARLIER APPLICATION NUMBER: 08/478,704
; EARLIER FILING DATE: 1995-06-07
```

```
; EARLIER APPLICATION NUMBER: 09/009,620
; EARLIER FILING DATE: 1998-01-20
; EARLIER APPLICATION NUMBER: 08/485,243
; EARLIER FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 09/008,172
; EARLIER FILING DATE: 1998-01-16
; EARLIER APPLICATION NUMBER: 08/482,711
; EARLIER FILING DATE: 1995-06-07
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 6
; LENGTH: 1430
; TYPE: PRT
; ORGANISM: streptococcus mutans
; US-09-210-361-6
```

```
Query Match          63.0%; Score 75; DB 3; Length 1430;
Best Local Similarity 71.4%; Pred. No. 0.0014;
Matches 15; Conservative 1; Mismatches 5; Indels 0; Gaps 0;
```

```
QY      1 ANDHSLILEAWSNDNTPYLHD 21
Db      495 AINHLILEAWSNDNDPQYKND 515
```

```
RESULT 9
; US-09-740-274-6
; Sequence 6, Application US/09740274
; Patent No. 6465203
; GENERAL INFORMATION:
; APPLICANT: Nichols, Scott E.
; TITLE OF INVENTION: Glucan-containing Compositions and Paper
; FILE REFERENCE: 0357CRD
; CURRENT APPLICATION NUMBER: US/09/740,274
; CURRENT FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: 09/210,361
; PRIOR FILING DATE: 1998-12-11
; PRIOR APPLICATION NUMBER: 09/007,999
; PRIOR FILING DATE: 1998-01-16
; PRIOR APPLICATION NUMBER: 08/478,704
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: 09/009,620
; PRIOR FILING DATE: 1998-01-20
; PRIOR APPLICATION NUMBER: 08/485,243
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: 09/008,172
; PRIOR FILING DATE: 1998-01-16
; PRIOR APPLICATION NUMBER: 08/482,711
; PRIOR FILING DATE: 1995-06-07
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 6
; LENGTH: 1430
; TYPE: PRT
; ORGANISM: streptococcus mutans
; US-09-740-274-6
```

```
Query Match          63.0%; Score 75; DB 4; Length 1430;
Best Local Similarity 71.4%; Pred. No. 0.0014;
Matches 15; Conservative 1; Mismatches 5; Indels 0; Gaps 0;
```

```
QY      1 ANDHSLILEAWSNDNTPYLHD 21
Db      495 AINHLILEAWSNDNDPQYKND 515
```

```
RESULT 10
; US-09-604-957-5
; Sequence 5, Application US/09604957
; Patent No. 6466314
; GENERAL INFORMATION:
; APPLICANT: VAN GEEL-SCHUTTEN, GERRITDINA HENDRIKA
; APPLICANT: DIJKHUIZEN, LUBBERT
```

APPLICANT: RAHAOU, HAKIM
APPLICANT: LEER, ROBERT-JAN
TITLE OF INVENTION: NOVEL GLUCOSYLTRANSFERASE AND NOVEL GLUCAN
FILE REFERENCE: BO 43388
CURRENT APPLICATION NUMBER: US/09/604,957
CURRENT FILING DATE: 2000-06-28
PRIOR APPLICATION NUMBER: 00201871.1
PRIOR FILING DATE: 2000-05-25
NUMBER OF SEQ ID NOS: 17
SOFTWARE: Patent In Ver. 2.1
SEQ ID NO 5
LENGTH: 523
TYPE: PRT
ORGANISM: Leuconostoc mesenteroides
US-09-604-957-5

Query Match 59.7%; Score 71; DB 4; Length 523;
Best Local Similarity 66.7%; Pred. No. 0.0018;
Matches 14; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 1 ANDHSLTEAMSDNDTPYLHD 21
DB 75 ANQHSLTEAMSDNDTPYLHD 95

RESULT 11
US-08-793-824-2
Sequence 2, Application US/08793824
Patent No. 5981838
GENERAL INFORMATION:
APPLICANT: Simpson, Christine Lynn
APPLICANT: Giffard, Philip Morrison
APPLICANT: Jacques, Nicholas Anthony
TITLE OF INVENTION: Genetic Manipulation of Plants to
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: Griffiths Hack & Co
STREET: Level 8, 168 Walker Street
CITY: No. 5981838th Sydney
STATE: New South Wales
COUNTRY: Australia
ZIP: 2060
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/793,824
FILING DATE:
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: AU PM7643
FILING DATE: 24-AUG-1994
TELECOMMUNICATION INFORMATION:
TELEPHONE: 61 2 957 5944
TELEFAX: 61 2 957 6288
TELEX: 26547
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1577 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: not relevant
MOLECULE TYPE: protein
ORIGINAL SOURCE:
ORGANISM: Streptococcus salivarius
US-08-793-824-2

Query Match 52.9%; Score 63; DB 2; Length 1577;
Best Local Similarity 72.2%; Pred. No. 0.14;
Matches 13; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 4 HSLTEAMSDNDTPYLHD 21
DB 591 HSLTEAMSDNDTPYLHD 608

RESULT 12
US-09-604-957-6
Sequence 6, Application US/09604957
Patent No. 6486314
GENERAL INFORMATION:
APPLICANT: VAN GEEL-SCHUTTEN, GERRITJINA HENDRIKA
APPLICANT: DIKHUIZEN, LUBBERT
APPLICANT: RAHAOU, HAKIM
APPLICANT: LEER, ROBERT-JAN
TITLE OF INVENTION: NOVEL GLUCOSYLTRANSFERASE AND NOVEL GLUCAN
FILE REFERENCE: BO 43388
CURRENT APPLICATION NUMBER: US/09/604,957
CURRENT FILING DATE: 2000-06-28
PRIOR APPLICATION NUMBER: 00201871.1
PRIOR FILING DATE: 2000-05-25
NUMBER OF SEQ ID NOS: 17
SOFTWARE: Patent In Ver. 2.1
SEQ ID NO 6
LENGTH: 584
TYPE: PRT
ORGANISM: Leuconostoc mesenteroides
US-09-604-957-6

Query Match 50.4%; Score 60; DB 4; Length 584;
Best Local Similarity 55.0%; Pred. No. 0.13;
Matches 11; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 1 ANDHSLTEAMSDNDTPYLHD 20
DB 75 ANKHSLTEAMSDNDTPYLHD 94

RESULT 13
US-09-499-203-2
Sequence 2, Application US/09499203
Patent No. 6570065
GENERAL INFORMATION:
APPLICANT: KOSSMANN, Jens
APPLICANT: WELSH, Thomas
APPLICANT: QUANZ, Martin
APPLICANT: KNUTH, Karola
TITLE OF INVENTION: Nucleic Acid Molecules Encoding Alternansucrase
FILE REFERENCE: 147-196P
CURRENT APPLICATION NUMBER: US/09/499,203
CURRENT FILING DATE: 2000-02-08
NUMBER OF SEQ ID NOS: 54
SOFTWARE: Patent In Ver. 2.1
SEQ ID NO 2
LENGTH: 2057
TYPE: PRT
ORGANISM: Leuconostoc mesenteroides
US-09-499-203-2

Query Match 50.4%; Score 60; DB 4; Length 2057;
Best Local Similarity 55.0%; Pred. No. 0.57;
Matches 11; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 1 ANDHSLTEAMSDNDTPYLHD 20
DB 665 ANKHSLTEAMSDNDTPYLHD 684

RESULT 14
US-09-604-957-7
Sequence 7, Application US/09604957
Patent No. 6486314
GENERAL INFORMATION:

```

? APPLICANT: VAN GEEL-SCHUTTEN, GERITDINA HENDRIKA
? APPLICANT: DICKHIZEN, LUBBERT
? APPLICANT: RAHOUI, HAKIM
? APPLICANT: LEER, ROBERT-JAN
? TITLE OF INVENTION: NOVEL GLUCOSYLTRANSFERASE AND NOVEL GLUCAN
? FILE REFERENCE: BO 43388
? CURRENT APPLICATION NUMBER: US/09/604,957
? CURRENT FILING DATE: 2000-06-28
? PRIOR APPLICATION NUMBER: 00201871.1
? PRIOR FILING DATE: 2000-05-25
? NUMBER OF SEQ ID NOS: 17
? SOFTWARE: Patentl Ver. 2.1
? SEQ ID NO 7
? LENGTH: 535
? TYPE: PR1
? ORGANISM: Lactobacillus reuteri
US-09-604-957-7

```

Query Match	41.2%	Score 49	DB 4	Length 535
Best Local Similarity	40.0%	Pred. No. 7		
Matches	8	Conservative	5	Mismatches 7; Indels 0; Gaps 0

```
QY      1 ANDHLSILEAWSNDNTPYLH 20  
       :|::||| |: |:  
Db     74 SNKHINILLEDMNHADPEYFN 93
```

RESULT 15
 US-09-604-957-3
 ; Sequence 3, Application US/09604957
 ; Patent No. 6486314
 ; GENERAL INFORMATION:
 ; APPLICANT: VAN GEL-SCHUTTEN, GERRITDIINA HENDRIKA
 ; APPLICANT: DIKHUIZEN, LUBBERT
 ; APPLICANT: RAHAOUI, HAKIM
 ; APPLICANT: LEER, ROBERT-JAN
 ; TITLE OF INVENTION: NOVEL GLUCOSYLTRANSFERASE AND NOVEL GLUCAN
 ; FILE REFERENCE: BO 43388
 ; CURRENT APPLICATION NUMBER: US/09/604,957
 ; CURRENT FILING DATE: 2000-06-28
 ; PRIOR APPLICATION NUMBER: 00201871.1
 ; PRIOR FILING DATE: 2000-05-25
 ; NUMBER OF SEQ ID NOS: 17
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 3
 ; LENGTH: 1278
 ; TYPE: PR1
 ; ORGANISM: Lactobacillus reuteri
 US-09-604-957-3

Query Match	41.2%	Score 49	DB 4	Length 1278
Best Local Similarity	40.0%	Pred. No. 20		
Matches	8	Mismatches	5	Indels 0
			7	Gaps 0

```
QY      1 ANDHLSILEAWSNDNTPYLH 20  
       :|::|||: |:  
Db     550 SNKHINILEDWNHADEFEYN 565
```

Search completed: November 13, 2003, 09:47:57
Job time : 15.6351 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 13, 2003, 09:31:40 ; Search time 38.5166 Seconds
(without alignments)
86.541 Million cell updates/sec

Title: US-09-290-049a-1
Perfect score: 119
Sequence: 1 ANDHSLRLAWSNDTPYLYLD 21

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues
Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: /SIDSI1/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
2: /SIDSI1/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
3: /SIDSI1/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
4: /SIDSI1/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
5: /SIDSI1/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
6: /SIDSI1/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
7: /SIDSI1/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
8: /SIDSI1/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
9: /SIDSI1/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
10: /SIDSI1/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
11: /SIDSI1/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
12: /SIDSI1/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
13: /SIDSI1/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
14: /SIDSI1/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
15: /SIDSI1/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
16: /SIDSI1/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
17: /SIDSI1/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
18: /SIDSI1/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
19: /SIDSI1/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
20: /SIDSI1/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
21: /SIDSI1/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
22: /SIDSI1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*
23: /SIDSI1/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*
24: /SIDSI1/gcgdata/geneseq/geneseq-emb1/AA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	119	100.0	1017	23	AAU79285
2	119	100.0	1475	23	AAU98027
3	119	100.0	1475	23	AAU98030
4	119	100.0	1475	23	AAU98031
5	119	100.0	1475	23	AAU98032
6	119	100.0	1475	23	AAU98033
7	119	100.0	1475	23	AAU98034
8	119	100.0	1475	23	AAU98035
9	119	100.0	1475	23	AAU98036

10	119	100.0	1475	23	AAU98037	S. mutans glucosyl
11	119	100.0	1475	23	AAU98038	S. mutans GTPB mut
12	119	100.0	1475	23	AAU98039	S. mutans glucosyl
13	119	100.0	1475	23	AAU98040	S. mutans glucosyl
14	119	100.0	1475	23	AAU98041	Streptococcus muta
15	110	92.4	1375	23	AAU98028	S. mutans glucosyl
16	110	92.4	1375	23	AAU98029	Streptococcus muta
17	108	90.8	1592	14	AAU98042	Streptococcus muta
18	75	63.0	1430	23	AAU98043	Glucosyltransferase
19	75	63.0	1430	23	AAU98044	S. mutans glucosyl
20	75	63.0	1430	23	AAU98045	S. mutans glucosyl
21	75	63.0	1430	23	AAU98046	S. mutans glucosyl
22	75	63.0	1430	23	AAU98047	S. mutans glucosyl
23	75	63.0	1430	23	AAU98048	S. mutans glucosyl
24	71	59.7	1527	23	AAU98049	S. mutans glucosyl
25	69	58.0	2835	23	AAU98050	Leuconostoc mesent
26	66	55.5	12	23	AAU98051	Dextran saccharase
27	63	52.9	1577	17	AAU98052	Dextran-saccharas
28	60	50.4	12	23	AAU98053	Alpha-D-glucosyltr
29	59	49.6	12	23	AAU98054	Dextran-saccharas
30	59	49.6	12	23	AAU98055	L. mesenteroides a
31	53	44.5	12	23	AAU98056	Dextran-saccharas
32	53	44.5	12	23	AAU98057	Dextran-saccharas
33	51	42.9	486	22	AAU98058	Dextran-saccharas
34	50	42.0	305	21	AAU98059	Lawsonia intracell
35	50	42.0	339	21	AAU98060	Arabidopsis thalia
36	50	42.0	649	22	AAU98061	Arabidopsis thalia
37	49	41.2	12	23	AAU98062	Proionibacterium
38	49	41.2	12	23	AAU98063	Dextran-saccharas
39	49	41.2	302	22	AAU98064	Salmonella typhi c
40	49	41.2	1781	23	AAU98065	Lactobacillus reut
41	48.5	40.8	401	22	AAU98066	Human polypeptide
42	48	40.3	855	23	AAU98067	Dextran saccharase
43	47.5	39.9	534	21	AAU98068	Arabidopsis aldehy
44	47	39.5	195	21	AAU98069	Arabidopsis thalia
45	47	39.5	338	15	AAU98070	T. niivum GAPDH.

ALIGNMENTS

RESULT 1
AAU79285
AAU79285 standard; Protein: 1017 AA.
AC AAU79285;
DT 13-AUG-2002 (first entry)
XX
DE Streptococcus mutans monoclonal antibody-related protein #2.
XX
KW Antibody; dental caries; water insoluble glucan synthetase;
KW anti-caries; glucosyl transferase-B; immunotherapy.
XX
OS Streptococcus mutans.
XX JP2002114709-A.
XX
PD 16-APR-2002.
XX
PF 04-OCT-2000; 2000UP-0304889.
XX
PR 04-OCT-2000; 2000UP-0304889.
XX
PA (UN1-) UNIV NIPPON.
XX WPI; 2002-448101/48.
XX
PT Anti-caries agent composed of a monoclonal antibody against an
PT inhibitory enzyme against water insoluble glucan synthetase of glucosyl
PT transferase-B (GTF-B) of Streptococcus mutans -
XX
PS Claim 4; Page 17-19; 28pp; Japanese.

XX The invention relates to a monoclonal antibody against dental caries and
 CC an anti-caries agent composed of a monoclonal antibody produced by
 CC Streptococcus mutans, particularly mouse hybridoma MHP126 (FERM P-17566)
 CC or mouse-hybridoma MHP136 (FERM P-17567), against an enzyme having
 CC inhibitive activity against water insoluble glucan synthetase of glucosyl
 CC transferase-B. The monoclonal antibody specifically inhibits water
 CC insoluble glucan synthetase of Streptococcus mutans produced glucosyl
 CC transferase-B and is used in the immunotherapy of dental caries. This
 CC sequence represents a Streptococcus mutans monoclonal antibody-related
 CC protein.
 CC
 XX Sequence 1017 AA;
 SQ
 Query Match 100.0%; Score 119; DB 23; Length 1017;
 Best Local Similarity 100.0%; Pred. No. 3,5e-09;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 1 ANDHLSILEAWSNDNTPYLHD 21
 447 ANDHLSILEAWSNDNTPYLHD 467
 OY
 RESULT 2
 AAU98027 standard; Protein; 1475 AA.
 ID AAU98027
 AC AAU98027;
 XX 27-AUG-2002 (first entry)
 DT
 DE S. mutans glucosyltransferase GTFB.
 XX
 KM Glucosyltransferase; GTFB; transgenic plant; paper sizing;
 XX coating composition; glucan; starch; latex; thermoplastic molecule;
 KM amyloplast; vacuole; paper manufacture.
 XX
 XX Streptococcus mutans.
 OS
 PN US2002031826-A1.
 XX
 PD 14-MAR-2002.
 XX
 PF 19-DEC-2000; 2000US-0740274.
 XX
 PR 11-DEC-1998; 98US-0210361.
 XX 07-JUN-1995; 95US-0478704.
 PR 07-JUN-1995; 95US-0482711.
 XX 07-JUN-1995; 95US-0485243.
 PR 16-JAN-1998; 98US-0007999.
 XX 16-JAN-1998; 98US-0008172.
 PR 20-JAN-1998; 98US-0009620.
 XX
 XX (NICH/) NICHOLS S E.
 PA
 PI
 XX
 PI
 XX
 PI
 XX
 DR WPI; 2002-414332/44.
 DR N-PDB; ABK52938.
 XX
 PT Glucosyltransferase B or D protein useful for producing a glucan useful
 PT as substitutes for and additions to modified starch and latexes in
 PT paper manufacture, comprises mutations in specific positions
 XX
 PS Disclosure; Page 21-25; 44pp; English.
 XX
 XX The invention an isolated protein comprising a glucosyltransferase
 CC (GTF) B polypeptide having changes at position from 1448V, D457N,
 CC D567T, K1014T, D457N/D567T, D457N/D571K, D567T/D571K,
 CC D567T/D571K/K1014T, 1448V/D457N/D567T/D571K/K779Q/K1014T,
 CC Y169A/Y170A/Y171A, and K779Q or a GTF D polypeptide having
 CC changes at positions from T589D, T589E, N471D, N471D/T589D, and
 CC N471D/T589E. Also included are a glucan produced by the GTF mutant,

CC an isolated polynucleotide which encodes P1 or P2, or its complementary
 CC polynucleotide, a ribonucleic acid sequence encoding the GTF mutant,
 CC an expression cassette comprising the polynucleotide operably linked to a
 CC promoter, a vector comprising the expression cassette, host cell
 CC introduced with the vector, a transgenic plant comprising the
 CC vector, a seed or tuber from the transgenic plant, a paper sizing and/or
 CC coating composition comprising a glucan produced in a plant transformed
 CC with a gene encoding the mutant GTF, wild type or, starch, a latex,
 CC thermoplastic molecule or their combinations or glucan and starch where
 CC the glucan is produced in the amyloplast and/or vacuole or a maize line
 CC deficient in starch biosynthesis, transformed with a gene encoding a
 CC glucosyltransferase B or D enzyme, wild-type or mutant and a paper
 CC comprising the glucan (paper sizing/coating agent). The vector is useful
 CC for producing a glucan in a plant. The method comprises transforming a
 CC plant cell with the vector, growing the plant cell under plant growing
 CC conditions to produce a regenerated plant and inducing expression of the
 CC polynucleotide for a time sufficient to produce the glucan in the
 CC regenerated plant, where the vector contains a transit sequence from
 CC ribulose biphosphate carboxylase small subunit, waxy, brittle-1 and
 CC chlorophyll AB binding protein to produce a transgenic plant, and glucan
 CC is produced in the amyloplast of potato or the vacuole of sugar beet.
 CC Glucans are useful as substitutes for and additions to modified starch
 CC and latexes in paper manufacture. Unlike prior art techniques, which
 CC require input materials that produce chemical effluents, paper
 CC manufacture utilizing the glucan produced by GTF, which utilizes
 CC biologically produced input materials, is more cost-effective and
 CC environmentally friendly. Moreover, glucans also exhibit thermoplastic
 CC properties and impart gloss to the paper during coating step.
 CC The present sequence represents GTFB.
 CC
 XX Sequence 1475 AA;
 SQ
 Query Match 100.0%; Score 119; DB 23; Length 1475;
 Best Local Similarity 100.0%; Pred. No. 5,4e-09;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 1 ANDHLSILEAWSNDNTPYLHD 21
 481 ANDHLSILEAWSNDNTPYLHD 501
 OY
 Db
 RESULT 3
 AAU98030
 ID AAU98030 standard; Protein; 1475 AA.
 XX
 AC AAU98030;
 XX
 DT 27-AUG-2002 (first entry)
 DE
 XX S. mutans glucosyltransferase GTFB mutant 1448V.
 XX
 KM Glucosyltransferase; GTFB; transgenic plant; paper sizing;
 KM coating composition; glucan; starch; latex; thermoplastic molecule;
 XX amyloplast; vacuole; paper manufacture; mutant; muten.
 XX
 XX Streptococcus mutans.
 OS
 PN Synthetic.
 XX
 PI
 XX
 PI
 XX
 PI
 XX
 DR Key Location/Qualifiers
 DR MISC-difference 448 /note= "Wild-type Ile substituted by Val"
 XX
 PT US2002031826-A1.
 XX
 PD 14-MAR-2002.
 XX
 PF 19-DEC-2000; 2000US-0740274.
 XX
 PR 11-DEC-1998; 98US-0210361.
 XX 07-JUN-1995; 95US-0478704.
 PR 07-JUN-1995; 95US-0482711.
 XX 07-JUN-1995; 95US-0485243.
 PR 16-JAN-1998; 98US-0007999.
 XX

PR 16-JAN-1998; 98US-0008172.
 PR 20-JAN-1998; 98US-0009620.
 PA (NICH/) NICHOLS S E.
 PI Nichols SE;
 XX WPI; 2002-414332/44.
 DR
 XX Glucosyltransferase B or D protein useful for producing a glucan useful
 PT as substitutes for and additions to modified starch and latexes in
 PT paper manufacture, comprises mutations in specific positions
 XX
 XX Claim 36; Page -; 44pp; English.
 XX
 CC The invention an isolated protein comprising a glucosyltransferase
 CC (GTF) B polypeptide having changes at position from I448V, D457N,
 CC D567T, K1014T, D457N/D567T, D457N/D571K, D567T/D571K,
 CC D567T/D571K/K1014T, I448V/D457N/D567T/D571K/K1014T,
 CC Y169A/Y170A/Y171A, and K779Q or a GTF D polypeptide having
 CC changes at positions from T589D, T589E, N471D, N471D/T589D, and
 CC N471D/T589E. Also included are a glucan produced by the GTF mutant,
 CC an isolated polynucleotide which encodes P1 or P2, or its complementary
 CC polynucleotide, a ribonucleic acid sequence encoding the GTF mutant,
 CC an expression cassette comprising the polynucleotide operably linked to a
 CC promoter, a vector comprising the expression cassette, host cell
 CC introduced with the vector, a transgenic plant comprising the
 CC vector, a seed or tuber from the transgenic plant, a paper sizing and/or
 CC coating composition comprising a glucan produced in a plant transformed
 CC with a gene encoding the mutant GTF, wild type or, starch, a latex,
 CC thermoplastic molecule or their combinations or glucan and starch where
 CC the glucan is produced in the amyloplast and/or vacuole or a maize line
 CC deficient in starch biosynthesis, transformed with a gene encoding a
 CC glucosyltransferase B or D enzyme, wild-type or mutant and a paper
 CC comprising the glucan (paper sizing/coating agent). The vector is useful
 CC for producing a glucan in a plant. The method comprises transforming a
 CC plant cell with the vector, growing the plant cell under plant growing
 CC conditions to produce a regenerated plant and inducing expression of the
 CC polynucleotide for a time sufficient to produce the glucan in the
 CC regenerated plant, where the vector contains a transit sequence from
 CC ribulose biphosphate carboxylase small subunit, waxy, brittle-1 and
 CC chlorophyll AB binding protein to produce a transgenic plant, and glucan
 CC is produced in the amyloplast of potato or the vacuole of sugar beet.
 CC Glucans are useful as substitutes for and additions to modified starch
 CC and latexes in paper manufacture. Unlike prior art techniques, which
 CC require input materials that produce chemical effluents, paper
 CC manufacture utilizing the glucan produced by GTF, which utilizes
 CC biologically produced input materials, is more cost-effective and
 CC environmentally friendly. Moreover, glucans also exhibit thermoplastic
 CC properties and impart gloss to the paper during coating step.
 CC The present sequence represents a GTFB mutant of the invention.
 CC Note: The present sequence is not shown in the specification but
 CC was created by the indexer using the GTFB sequence appearing as AAU98027
 CC and the information in claim 36.
 XX
 SQ Sequence 1475 AA;
 Query March 100.0%; Score 119; DB 23; Length 1475;
 Best Local Similarity 100.0%; Pred. NO.5.4e-09;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 ANDHSLILEMSDNDPTPIHD 21
 Db 481 ANDHSLILEMSDNDPTPIHD 501
 RESULT 4
 AAU98031
 ID AAU98031 standard; Protein; 1475 AA.
 XX
 AC AAU98031;
 XX
 DT 27-AUG-2002 (first entry)

XX
 DE S. mutans glucosyltransferase GTFB mutant D457N.
 XX
 KW Glucosyltransferase; GTFB; transgenic plant; paper sizing;
 KW coating composition; glucan; starch; latex; thermoplastic molecule;
 KW amyloplast; vacuole; paper manufacture; mutant; muten.
 XX
 OS Streptococcus mutans.
 OS Synthetic.
 XX
 FH Key Location/Qualifiers
 FT Msc-difference 457
 FT /note="Wild-type Asp substituted by Asn"
 XX
 PN US2002031826-A1.
 PD 14-MAR-2002.
 XX
 PF 19-DEC-2000; 2000US-0740274.
 XX
 PR 11-DEC-1998; 98US-0210361.
 PR 07-JUN-1995; 95US-0478704.
 PR 07-JUN-1995; 95US-0482711.
 PR 07-JUN-1995; 95US-0485243.
 PR 16-JAN-1998; 98US-0007999.
 PR 16-JAN-1998; 98US-0008172.
 PR 20-JAN-1998; 98US-0009620.
 XX
 PA (NICH/) NICHOLS S E.
 PI Nichols SE;
 XX WPI; 2002-414332/44.
 DR
 XX Glucosyltransferase B or D protein useful for producing a glucan useful
 PT as substitutes for and additions to modified starch and latexes in
 PT paper manufacture, comprises mutations in specific positions
 XX
 XX Claim 36; Page -; 44pp; English.
 XX
 CC The invention an isolated protein comprising a glucosyltransferase
 CC (GTF) B polypeptide having changes at position from I448V, D457N,
 CC D567T, K1014T, D457N/D567T, D457N/D571K, D567T/D571K,
 CC D567T/D571K/K1014T, I448V/D457N/D567T/D571K/K1014T,
 CC Y169A/Y170A/Y171A, and K779Q or a GTF D polypeptide having
 CC changes at positions from T589D, T589E, N471D, N471D/T589D, and
 CC N471D/T589E. Also included are a glucan produced by the GTF mutant,
 CC an isolated polynucleotide which encodes P1 or P2, or its complementary
 CC polynucleotide, a ribonucleic acid sequence encoding the GTF mutant,
 CC an expression cassette comprising the polynucleotide operably linked to a
 CC promoter, a vector comprising the expression cassette, host cell
 CC introduced with the vector, a transgenic plant comprising the
 CC vector, a seed or tuber from the transgenic plant, a paper sizing and/or
 CC coating composition comprising a glucan produced in a plant transformed
 CC with a gene encoding the mutant GTF, wild type or, starch, a latex,
 CC thermoplastic molecule or their combinations or glucan and starch where
 CC the glucan is produced in the amyloplast and/or vacuole or a maize line
 CC deficient in starch biosynthesis, transformed with a gene encoding a
 CC glucosyltransferase B or D enzyme, wild-type or mutant and a paper
 CC comprising the glucan (paper sizing/coating agent). The vector is useful
 CC for producing a glucan in a plant. The method comprises transforming a
 CC plant cell with the vector, growing the plant cell under plant growing
 CC conditions to produce a regenerated plant and inducing expression of the
 CC polynucleotide for a time sufficient to produce the glucan in the
 CC regenerated plant, where the vector contains a transit sequence from
 CC ribulose biphosphate carboxylase small subunit, waxy, brittle-1 and
 CC chlorophyll AB binding protein to produce a transgenic plant, and glucan
 CC is produced in the amyloplast of potato or the vacuole of sugar beet.
 CC Glucans are useful as substitutes for and additions to modified starch
 CC and latexes in paper manufacture. Unlike prior art techniques, which
 CC require input materials that produce chemical effluents, paper
 CC manufacture utilizing the glucan produced by GTF, which utilizes
 CC biologically produced input materials, is more cost-effective and

environmentally friendly. Moreover, glucans also exhibit thermoplastic properties and impart gloss to the paper during coating step.

The present sequence represents a GTFB mutant of the invention.

Note: The present sequence is not shown in the specification but was created by the indexer using the GTFB sequence appearing as AAU98027 and the information in claim 36.

Sequence 1475 AA;

Query Match 100.0%; Score 119; DB 23; Length 1475;
Best Local Similarity 100.0%; Pred. No. 5,4e-09;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 ANDHSTILEAWSNDPTYLHD 21
481 ANDHSTILEAWSNDPTYLHD 501

RESULT 5
AAU98032
ID AAU98032 standard; Protein; 1475 AA.

AAU98032;
27-AUG-2002 (first entry)

S. mutans glucosyltransferase GTFB mutant D567T.

Glucosyltransferase; GTFB; transgenic plant; paper sizing;
coating composition; glucan; starch; latex; thermoplastic molecule;
amyloplast; vacuole; paper manufacture; mutant; mutein.

Streptococcus mutans.
Synthetic.

Key Location/Qualifiers
Misc-difference 567 /note="Wild-type Asp substituted by Thr"

US2002031826-A1.

14-MAR-2002.

19-DEC-2000; 2000US-0740274.

11-DEC-1998; 98US-0210361.
07-JUN-1995; 95US-0478704.
07-JUN-1995; 95US-0482711.
07-JUN-1995; 95US-0485243.
16-JAN-1998; 98US-0007999.
16-JAN-1998; 98US-0008172.
20-JAN-1998; 98US-0009620.

(NICH/) NICHOLS S E.
Nichols SE;
MPI; 2002-414332/44.

Glucosyltransferase B or D protein useful for producing a glucan useful as substitutes for and additions to modified starch and latexes in paper manufacture, comprises mutations in specific positions -

Claim 36; Page -; 44pp; English.

The invention an isolated protein comprising a glucosyltransferase (GTF) B polypeptide having changes at position from I448V, D457N, D567T, K1014T, D457N/D567T, D457N/D571K, D567T/D571K, D567T/D571K/K1014T, I448V/D457N/D567T/D571K/K779G/K1014T, Y169A/Y170A/Y171A, and K779Q or a GTF D polypeptide having changes at positions from T589D, T589E, N471D, N471D/T589D, and N471D/T589E. Also included are a glucan produced by the GTF mutant, an isolated polynucleotide which encodes P1 or P2, or its complementary

polynucleotide, a ribonucleic acid sequence encoding the GTF mutant, an expression cassette comprising the polynucleotide operably linked to a promoter, a vector comprising the expression cassette, host cell introduced with the vector, a transgenic plant comprising the vector, a seed or tuber from the transgenic plant, a paper sizing and/or coating composition comprising a glucan produced in a plant transformed with a gene encoding the mutant GTF, wild type or, starch, a latex, thermoplastic molecule or their combinations or glucan and starch where the glucan is produced in the amyloplast and/or vacuole or a maize line deficient in starch biosynthesis, transformed with a gene encoding a glucosyltransferase B or D enzyme, wild-type or mutant and a paper comprising the glucan (paper sizing/coating agent). The vector is useful for producing a glucan in a plant. The method comprises transforming a plant cell with the vector, growing the plant cell under plant growing conditions to produce a regenerated plant and inducing expression of the polynucleotide for a time sufficient to produce the glucan in the regenerated plant, where the vector contains a transit sequence from ribulose biphosphate carboxylase small subunit, waxy, brittle-1 and chlorophyll AB binding protein of potato or the vacuole of sugar beet.

Glucans are useful as substitutes for and additions to modified starch and latexes in paper manufacture. Unlike prior art techniques, which require input materials that produce chemical effluents, paper manufacture utilizing the glucan produced by GTF, which utilizes biologically produced input materials, is more cost-effective and environmentally friendly. Moreover, glucans also exhibit thermoplastic properties and impart gloss to the paper during coating step.

Note: The present sequence represents a GTFB mutant of the invention but was created by the indexer using the GTFB sequence appearing as AAU98027 and the information in claim 36.

Sequence 1475 AA;

Query Match 100.0%; Score 119; DB 23; Length 1475;
Best Local Similarity 100.0%; Pred. No. 5,4e-09;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 ANDHSTILEAWSNDPTYLHD 21
481 ANDHSTILEAWSNDPTYLHD 501

RESULT 6
AAU98033
ID AAU98033 standard; Protein; 1475 AA.

AAU98033;
27-AUG-2002 (first entry)

S. mutans glucosyltransferase GTFB mutant K1014T.

Glucosyltransferase; GTFB; transgenic plant; paper sizing;
coating composition; glucan; starch; latex; thermoplastic molecule;
amyloplast; vacuole; paper manufacture; mutant; mutein.

Streptococcus mutans.
Synthetic.

Key Location/Qualifiers
Misc-difference 1014 /note="Wild-type Lys substituted by Thr"

US2002031826-A1.

14-MAR-2002.

19-DEC-2000; 2000US-0740274.

11-DEC-1998; 98US-0210361.
07-JUN-1995; 95US-0478704.
07-JUN-1995; 95US-0482711.

PR 07-JUN-1995; 95US-0485243.
 PR 16-JAN-1998; 98US-0007999.
 PR 16-JAN-1998; 98US-0008172.
 PR 20-JAN-1998; 98US-0009620.
 XX (NICH/) NICHOLS S E.
 PA Nichols SE;
 PI WPI; 2002-414332/44.
 DR WPI; 2002-414332/44.
 XX Glucosyltransferase B or D protein useful for producing a glucan useful
 PT as substitutes for and additions to modified starch and latexes in
 PT paper manufacture, comprises mutations in specific positions -
 PS Claim 36; Page -; 44pp; English.
 XX The invention an isolated protein comprising a glucosyltransferase
 CC (GTF) B polypeptide having changes at position from 1448V, D457N,
 CC D567T, K1014T, D457N/D567T, D457N/D571K, D567T/D571K,
 CC D567T/D571K/K1014T, 1448V/D457N/D567T/D571K/K779Q/K1014T,
 CC Y169A/Y170A/Y171A, and K779Q or a GTF D polypeptide having
 CC changes at positions from T589D, T589E, N471D, N471D/T589D, and
 CC N471D/T589E. Also included are a glucan produced by the GTF mutant,
 CC an isolated polynucleotide which encodes P1 or P2, or its complementary
 CC polynucleotide, a ribonucleic acid sequence encoding the GTF mutant,
 CC an expression cassette comprising the polynucleotide operably linked to a
 CC promoter, a vector comprising the expression cassette, host cell
 CC introduced with the vector, a transgenic plant comprising the
 CC vector, a seed or tuber from the transgenic plant, a paper sizing and/or
 CC coating composition comprising a glucan produced in a plant transformed
 CC with a gene encoding the mutant GTF, wild type or, starch, a latex,
 CC thermoplastic molecule or their combinations or glucan and starch where
 CC the glucan is produced in the amyloplast and/or vacuole or a maize line
 CC deficient in starch biosynthesis, transformed with a gene encoding a
 CC glucosyltransferase B or D enzyme, wild-type or mutant and a paper
 CC comprising the glucan (paper sizing/coating agent). The vector is useful
 CC for producing a glucan in a plant. The method comprises transforming a
 CC plant cell with the vector, growing the plant cell under plant growing
 CC conditions to produce a regenerated plant and inducing expression of the
 CC polynucleotide for a time sufficient to produce the glucan in the
 CC regenerated plant, where the vector contains a transit sequence from
 CC ribulose biphosphate carboxylase small subunit, waxy, brittle-1 and
 CC chlorophyll AB binding protein to produce a transgenic plant, and glucan
 CC is produced in the amyloplast of potato or the vacuole of sugar beet.
 CC Glucans are useful as substitutes for and additions to modified starch
 CC and latexes in paper manufacture. Unlike prior art techniques, which
 CC require input materials that produce chemical effluents, paper
 CC manufacture utilizing the glucan produced by GTF, which utilizes
 CC biologically produced input materials, is more cost-effective and
 CC environmentally friendly. Moreover, glucans also exhibit thermoplastic
 CC properties and impart gloss to the paper during coating step.
 CC The present sequence represents a GTFB mutant of the invention.
 CC Note: The present sequence is not shown in the specification but
 CC was created by the indexer using the GTFB sequence appearing as AAU98027
 CC and the information in claim 36.
 XX Sequence 1475 AA;
 SQ Query Match 100.0%; Score 119; DB 23; Length 1475;
 Best Local Similarity 100.0%; Pred. No. 5.4e-09;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 ANDHSLTEAMSDNDPTFLND 21
 Db 481 ANDHSLTEAMSDNDPTFLND 501
 RESULT 7
 AAU98034
 ID AAU98034 standard; Protein; 1475 AA.
 XX
 AC AAU98034;

XX 27-AUG-2002 (first entry)
 DT S. mutans glucosyltransferase GTFB mutant D457N/D567T.
 XX
 DE Glucosyltransferase; GTFB; transgenic plant; paper sizing;
 XX Glucan; starch; latex; thermoplastic molecule;
 KW amyloplast; vacuole; paper manufacture; mutant; mutin.
 XX
 OS Streptococcus mutans.
 OS Synthetic.
 XX
 FH Key Location/Qualifiers
 FT Misc-difference 457
 FT /note= "Wild-type Asp substituted by Asn"
 FT Misc-difference 567
 FT /note= "Wild-type Asp substituted by Thr"
 XX
 XX US2002031826-A1.
 XX
 PD 14-MAR-2002.
 XX
 XX 19-DEC-2000; 2000US-0740274.
 XX
 XX 11-DEC-1998; 98US-0210361.
 XX
 PR 07-JUN-1995; 95US-0478704.
 PR 07-JUN-1995; 95US-0482711.
 PR 07-JUN-1995; 95US-0485243.
 PR 16-JAN-1998; 98US-0007999.
 PR 16-JAN-1998; 98US-0008172.
 PR 20-JAN-1998; 98US-0009620.
 XX
 PA (NICH/) NICHOLS S E.
 PI Nichols SE;
 DR WPI; 2002-414332/44.
 XX Glucosyltransferase B or D protein useful for producing a glucan useful
 PT as substitutes for and additions to modified starch and latexes in
 PT paper manufacture, comprises mutations in specific positions -
 PS Claim 36; Page -; 44pp; English.
 XX The invention an isolated protein comprising a glucosyltransferase
 CC (GTF) B polypeptide having changes at position from 1448V, D457N,
 CC D567T, K1014T, D457N/D567T, D457N/D571K, D567T/D571K,
 CC D567T/D571K/K1014T, 1448V/D457N/D567T/D571K/K779Q/K1014T,
 CC Y169A/Y170A/Y171A, and K779Q or a GTF D polypeptide having
 CC changes at positions from T589D, T589E, N471D, N471D/T589D, and
 CC N471D/T589E. Also included are a glucan produced by the GTF mutant,
 CC an isolated polynucleotide which encodes P1 or P2, or its complementary
 CC polynucleotide, a ribonucleic acid sequence encoding the GTF mutant,
 CC an expression cassette comprising the polynucleotide operably linked to a
 CC promoter, a vector comprising the expression cassette, host cell
 CC introduced with the vector, a transgenic plant comprising the
 CC vector, a seed or tuber from the transgenic plant, a paper sizing and/or
 CC coating composition comprising a glucan produced in a plant transformed
 CC with a gene encoding the mutant GTF, wild type or, starch, a latex,
 CC thermoplastic molecule or their combinations or glucan and starch where
 CC the glucan is produced in the amyloplast and/or vacuole or a maize line
 CC deficient in starch biosynthesis, transformed with a gene encoding a
 CC glucosyltransferase B or D enzyme, wild-type or mutant and a paper
 CC comprising the glucan (paper sizing/coating agent). The vector is useful
 CC for producing a glucan in a plant. The method comprises transforming a
 CC plant cell with the vector, growing the plant cell under plant growing
 CC conditions to produce a regenerated plant and inducing expression of the
 CC polynucleotide for a time sufficient to produce the glucan in the
 CC regenerated plant, where the vector contains a transit sequence from
 CC ribulose biphosphate carboxylase small subunit, waxy, brittle-1 and
 CC chlorophyll AB binding protein to produce a transgenic plant, and glucan
 CC is produced in the amyloplast of potato or the vacuole of sugar beet.
 CC Glucans are useful as substitutes for and additions to modified starch

CC and latexes in paper manufacture. Unlike prior art techniques, which
 CC require input materials that produce chemical effluents, paper
 CC manufacture utilizing the glucan produced by GTF, which utilizes
 CC biologically produced input materials, is more cost-effective and
 CC environmentally friendly. Moreover, glucans also exhibit thermoplastic
 CC properties and impart gloss to the paper during coating step.
 CC The present sequence represents a GTFB mutant of the invention.
 CC Note: The present sequence is not shown in the specification but
 CC was created by the indexer using the GTFB sequence appearing as AAU98027
 CC and the information in claim 36.

XX SQ Sequence 1475 AA;

Query Match 100.0%; Score 119; DB 23; Length 1475;
 Best Local Similarity 100.0%; Pred. No. 5.4e-09;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ANDHLSILEAWSNDPTPLHD 21
 |||||
 DB 481 ANDHLSILEAWSNDPTPLHD 501

RESULT 9
 AAU98035 AAU98035 standard; Protein; 1475 AA.

XX AC AAU98035;

XX DT 27-AUG-2002 (first entry)

XX S. mutans glucosyltransferase GTFB mutant D457N/D571K.

XX Glucosyltransferase; GTFB; transgenic plant; paper sizing;
 XX coating composition; glucan; starch; latex; thermoplastic molecule;
 XX amyloplast; vacuole; paper manufacture; mutant; muten.

XX Streptococcus mutans.

OS Synthetic.

XX Key Location/Qualifiers

FT Misc-difference 457 /note= "Wild-type Asp substituted by Asn"
 FT Misc-difference 571 /note= "Wild-type Asp substituted by Lys"

XX US2002031826-A1.
 XX 14-MAR-2002.

XX 19-DEC-2000; 2000US-0740274.
 XX 11-DEC-1998; 98US-0210361.
 XX 07-JUN-1995; 95US-0478704.
 XX 07-JUN-1995; 95US-0482711.
 XX 07-JUN-1995; 95US-0485243.
 XX 16-JAN-1998; 98US-0007999.
 XX 16-JAN-1998; 98US-0008172.
 XX 20-JAN-1998; 98US-0009620.

XX (NICH/) NICHOLS S E.

XX Nichols SE;

XX WPI; 2002-414332/44.

XX Glucosyltransferase B or D protein useful for producing a glucan useful
 PT as substitutes for and additions to modified starch and latexes in
 PT paper manufacture, comprises mutations in specific positions -
 XX Claim 36; Page -; 44p; English.

CC The invention an isolated protein comprising a glucosyltransferase
 CC (GTF) B polypeptide having changes at position from I448V, D457N,

CC D567T, K1014T, D457N/D567T, D457N/D571K, D567T/D571K.

CC D567T/D571K/K1014T, 1448V/D457N/D567T/D571K/K779G/K1014T,

CC Y169A/Y170A/Y171A, and K779G or a GTF D polypeptide having

CC changes at positions from T589D, T589E, N471D, N471D/T589D, and

CC N471D/T589E. Also included are a glucan produced by the GTF mutant,

CC an isolated polynucleotide which encodes P1 or P2, or its complementary

CC polynucleotide, a ribonucleic acid sequence encoding the GTF mutant,

CC an expression cassette comprising the polynucleotide operably linked to a

CC promoter, a vector comprising the expression cassette, host cell

CC introduced with the vector, a transgenic plant comprising the

CC vector, a seed or tuber from the transgenic plant, a paper sizing and/or

CC coating composition comprising a glucan produced in a plant transformed

CC with a gene encoding the mutant GTF, wild type or, starch, a latex.

CC thermoplastic molecule or their combinations or glucan and starch where

CC the glucan is produced in the amyloplast and/or vacuole or a maize line

CC deficient in starch biosynthesis, transformed with a gene encoding a

CC glucosyltransferase B or D enzyme, wild-type or mutant and a paper

CC comprising the glucan (paper sizing/coating agent). The vector is useful

CC for producing a glucan in a plant. The method comprises transforming a

CC plant cell with the vector, growing the plant cell under plant growing

CC conditions to produce a regenerated plant and inducing expression of the

CC polynucleotide for a time sufficient to produce the glucan in the

CC regenerated plant, where the vector contains a transit sequence from

CC ribulose biphosphate carboxylase small subunit, waxy, brittle-1 and

CC chlorophyll AB binding protein to produce a transgenic plant, and glucan

CC is produced in the amyloplast of potato or and additions to modified starch

CC Glucans are useful as substitutes for and additions to modified starch

CC and latexes in paper manufacture. Unlike prior art techniques, which

CC require input materials that produce chemical effluents, paper

CC manufacture utilizing the glucan produced by GTF, which utilizes

CC biologically produced input materials, is more cost-effective and

CC environmentally friendly. Moreover, glucans also exhibit thermoplastic

CC properties and impart gloss to the paper during coating step.

CC The present sequence represents a GTFB mutant of the invention.

CC Note: The present sequence is not shown in the specification but

CC was created by the indexer using the GTFB sequence appearing as AAU98027

CC and the information in claim 36.

XX SQ Sequence 1475 AA;

Query Match 100.0%; Score 119; DB 23; Length 1475;
 Best Local Similarity 100.0%; Pred. No. 5.4e-09;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ANDHLSILEAWSNDPTPLHD 21
 |||||
 DB 481 ANDHLSILEAWSNDPTPLHD 501

RESULT 9
 AAU98036 AAU98036 standard; Protein; 1475 AA.

XX AC AAU98036;

XX DT 27-AUG-2002 (first entry)

XX S. mutans glucosyltransferase GTFB mutant D567T/D571K.

XX Glucosyltransferase; GTFB; transgenic plant; paper sizing;
 XX coating composition; glucan; starch; latex; thermoplastic molecule;
 XX amyloplast; vacuole; paper manufacture; mutant; muten.

XX Streptococcus mutans.

OS Synthetic.

XX Key Location/Qualifiers

FT Misc-difference 567 /note= "Wild-type Asp substituted by Thr"
 FT Misc-difference 571 /note= "Wild-type Asp substituted by Lys"

XX US2002031826-A1.

XX 14-MAR-2002.
 PD 19-DEC-2000; 2000US-0740274.
 XX
 XX
 PR 11-DEC-1998; 98US-0210361.
 XX 07-JUN-1995; 95US-0478704.
 PR 07-JUN-1995; 95US-0482711.
 PR 07-JUN-1995; 95US-0485243.
 PR 16-JAN-1998; 98US-0007999.
 PR 16-JAN-1998; 98US-0008172.
 PR 20-JAN-1998; 98US-0009620.
 XX
 XX (NICH/) NICHOLS S E.
 PA
 XX
 PI Nichols SE;
 XX
 XX WPI; 2002-414332/44.
 DR
 XX
 XX
 PT Glucosyltransferase B or D protein useful for producing a glucan useful
 PT as substitutes for and additions to modified starch and latexes in
 PT paper manufacture, comprises mutations in specific positions -
 XX
 XX Claim 36; Page -; 44pp; English.

XX The invention an isolated protein comprising a glucosyltransferase
 CC (GTF) B polypeptide having changes at position from 1448V, D457N,
 CC D567T, K1014T, D457N/D567T, D457N/D571K, D567T/D571K,
 CC D567T/D571K/K1014T, 1448V/D457N/D567T/D571K/K779Q/K1014T,
 CC Y169A/Y170A/Y171A, and K779Q or a GTF D polypeptide having
 CC changes at positions from T589D, T589E, N471D, N471D/T589D, and
 CC N471D/T589E. Also included are a glucan produced by the GTF mutant,
 CC an isolated polynucleotide which encodes P1 or P2, or its complementary
 CC polynucleotide, a ribonucleic acid sequence encoding the GTF mutant,
 CC an expression cassette comprising the polynucleotide operably linked to a
 CC promoter, a vector comprising the expression cassette, host cell
 CC introduced with the vector, a transgenic plant comprising the
 CC vector, a seed or tuber from the transgenic plant, a paper string and/or
 CC coating composition comprising a glucan produced in a plant transformed
 CC with a gene encoding the mutant GTF, wild type or, starch, a latex,
 CC thermoplastic molecule or their combinations or glucan and starch where
 CC the glucan is produced in the amyloplast and/or vacuole or a maize line
 CC deficient in starch biosynthesis, transformed with a gene encoding a
 CC glucosyltransferase B or D enzyme, wild-type or mutant and a paper
 CC comprising the glucan (paper sizing/coating agent). The vector is useful
 CC for producing a glucan in a plant. The method comprises transforming a
 CC plant cell with the vector, growing the plant cell under plant growing
 CC conditions to produce a regenerated plant and inducing expression of the
 CC polynucleotide for a time sufficient to produce the glucan in the
 CC regenerated plant, where the vector contains a transit sequence from
 CC ribulose biphosphate carboxylase small subunit, waxy, brittle-1 and
 CC chlorophyll AB binding protein to produce a transgenic plant, and glucan
 CC is produced in the amyloplast of potato or the vacuole of sugar beet.
 CC Glucans are useful as substitutes for and additions to modified starch
 CC and latexes in paper manufacture. Unlike prior art techniques, which
 CC require input materials that produce chemical effluents, paper
 CC manufacture utilizing the glucan produced by GTF, which utilizes
 CC biologically produced input materials, is more cost-effective and
 CC environmentally friendly. Moreover, glucans also exhibit thermoplastic
 CC properties and impart gloss to the paper during coating step.
 CC The present sequence represents a GTFB mutant of the invention.
 CC Note: The present sequence is not shown in the specification but
 CC was created by the indexer using the GTFB sequence appearing as AAU98027
 CC and the information in claim 36.

XX Sequence 1475 AA;

Query Match 100.0%; Score 119; DB 23; Length 1475;
 Best Local Similarity 100.0%; Pred. No. 5,4e-09;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ANDHSLTEAMSNDPTPLHD 21
 |||||||

DB 461 ANDHSLTEAMSNDPTPLHD 501
 XX
 XX RESULT 10
 XX AAU98037
 ID AAU98037 standard; Protein; 1475 AA.
 XX
 XX AAU98037;
 XX
 XX 27-AUG-2002 (first entry)
 XX
 DE S. mutans glucosyltransferase GTFB mutant D567T/D571K/K1014T.
 XX
 XX Glucosyltransferase; GTFB; transgenic plant; paper string;
 XX coating composition; glucan; starch; latex; thermoplastic molecule;
 XX amyloplast; vacuole; paper manufacture; mutant; mutain.
 XX
 OS Streptococcus mutans.
 OS Synthetic.
 XX
 XX Key Location/Qualifiers
 XX
 XX Misc-difference 567
 XX /note= "Wild-type Asp substituted by Thr"
 XX
 XX Misc-difference 571
 XX /note= "Wild-type Asp substituted by Lys"
 XX
 XX Misc-difference 1014
 XX /note= "Wild-type Lys substituted by Thr"
 XX
 XX US2002031826-A1.
 XX
 PD 14-MAR-2002.
 XX
 XX
 PF 19-DEC-2000; 2000US-0740274.
 XX
 XX
 PR 11-DEC-1998; 98US-0210361.
 PR 07-JUN-1995; 95US-0478704.
 PR 07-JUN-1995; 95US-0482711.
 PR 07-JUN-1995; 95US-0485243.
 PR 16-JAN-1998; 98US-0007999.
 PR 16-JAN-1998; 98US-0008172.
 PR 20-JAN-1998; 98US-0009620.
 XX
 XX (NICH/) NICHOLS S E.
 XX
 XX Nichols SE;
 XX
 XX WPI; 2002-414332/44.
 DR
 XX
 XX
 PT Glucosyltransferase B or D protein useful for producing a glucan useful
 PT as substitutes for and additions to modified starch and latexes in
 PT paper manufacture, comprises mutations in specific positions -
 XX
 XX Claim 36; Page -; 44pp; English.

XX The invention an isolated protein comprising a glucosyltransferase
 CC (GTF) B polypeptide having changes at position from 1448V, D457N,
 CC D567T, K1014T, D457N/D567T, D457N/D571K, D567T/D571K,
 CC D567T/D571K/K1014T, 1448V/D457N/D567T/D571K/K779Q/K1014T,
 CC Y169A/Y170A/Y171A, and K779Q or a GTF D polypeptide having
 CC changes at positions from T589D, T589E, N471D, N471D/T589D, and
 CC N471D/T589E. Also included are a glucan produced by the GTF mutant,
 CC an isolated polynucleotide which encodes P1 or P2, or its complementary
 CC polynucleotide, a ribonucleic acid sequence encoding the GTF mutant,
 CC an expression cassette comprising the polynucleotide operably linked to a
 CC promoter, a vector comprising the expression cassette, host cell
 CC introduced with the vector, a transgenic plant comprising the
 CC vector, a seed or tuber from the transgenic plant, a paper string and/or
 CC coating composition comprising a glucan produced in a plant transformed
 CC with a gene encoding the mutant GTF, wild type or, starch, a latex,
 CC thermoplastic molecule or their combinations or glucan and starch where
 CC the glucan is produced in the amyloplast and/or vacuole or a maize line
 CC deficient in starch biosynthesis, transformed with a gene encoding a
 CC glucosyltransferase B or D enzyme, wild-type or mutant and a paper

CC comprising the glucan (paper sizing/coating agent). The vector is useful
 CC for producing a glucan in a plant. The method comprises transforming a
 CC plant cell with the vector, growing the plant cell under plant growing
 CC conditions to produce a regenerated plant and inducing expression of the
 CC polynucleotide for a time sufficient to produce the glucan in the
 CC regenerated plant, where the vector contains a transit sequence from
 CC ribulose biphosphate carboxylase small subunit, waxy, brittle-1 and
 CC chlorophyll AB binding protein to produce a transgenic plant, and glucan
 CC is produced in the amyloplast of potato or the vacuole of sugar beet.
 CC Glucans are useful as substitutes for and additions to modified starch
 CC and latexes in paper manufacture. Unlike prior art techniques, which
 CC require input materials that produce chemical effluents, paper
 CC manufacture utilizing the glucan produced by GFP, which utilizes
 CC biologically produced input materials, is more cost-effective and
 CC environmentally friendly. Moreover, glucans also exhibit thermoplastic
 CC properties and impart gloss to the paper during coating step.
 CC The present sequence represents a GFP mutant of the invention.
 CC Note: The present sequence is not shown in the specification but
 CC was created by the indexer using the GFPB sequence appearing as AAU98027
 CC and the information in claim 36.

XX Sequence 1475 AA;
 SQ Query Match 100.0%; Score 119; DB 23; Length 1475;
 Best Local Similarity 100.0%; Pred. No. 5.4e-09;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ANDHSLILEAWSNDNTPYLHD 21
 Db 481 ANDHSLILEAWSNDNTPYLHD 501

RESULT 11
 AAU98038 ID AAU98038 standard; Protein; 1475 AA.
 XX AC AAU98038;
 XX DT 27-AUG-2002 (first entry)
 DE 5. mutants GFPB mutant 1448V/D457N/D567T/D571K/K779Q/K1014T.
 XX XX Glucosyltransferase; GFPB; transgenic plant; paper sizing;
 KM coating composition; glucan; starch; latex; thermoplastic molecule;
 KM amyloplast; vacuole; paper manufacture; mutant; mutain.
 XX OS Streptococcus mutans.
 OS Synthetic.
 XX FH Key Location/Qualifiers
 FT Misc-difference 448 /note= "Wild-type Ile substituted by Val"
 FT FT Misc-difference 457 /note= "Wild-type Asp substituted by Asn"
 FT FT Misc-difference 567 /note= "Wild-type Asp substituted by Thr"
 FT FT Misc-difference 571 /note= "Wild-type Asp substituted by Lys"
 FT FT Misc-difference 779 /note= "Wild-type Lys substituted by Gln"
 FT FT Misc-difference 1014 /note= "Wild-type Lys substituted by Thr"
 FT XX US2002031826-A1.
 XX PN 14-MAR-2002.
 XX PD 19-DEC-2000; 2000US-0740274.
 XX PF 11-DEC-1998; 98US-0210361.
 XX PR 07-JUN-1995; 95US-0478704.
 XX PR 07-JUN-1995; 95US-0482711.
 XX PR 07-JUN-1995; 95US-0485243.

PR 16-JAN-1998; 98US-0007999.
 PR 16-JAN-1998; 98US-0008172.
 PR 20-JAN-1998; 98US-0009620.
 XX (NICH/) NICHOLS S E.
 XX Nichols SE;
 XX WPI; 2002-414332/44.
 XX DR
 XX PT Glucosyltransferase B or D protein useful for producing a glucan useful
 PT as substitutes for and additions to modified starch and latexes in
 PT paper manufacture, comprises mutations in specific positions
 PS Claim 36, Page -; 44pp; English.

CC The invention an isolated protein comprising a glucosyltransferase
 CC (GTF) B polypeptide having changes at position from 1448V, D457N,
 CC D567T, K1014T, D457N/D567T, D457N/D571K, D567T/D571K,
 CC D567T/D571K/K1014T, 1448V/D457N/D567T/D571K/K779Q/K1014T,
 CC Y169A/Y170A/Y171A, and K779Q or a GFP D polypeptide having
 CC changes at positions from T589D, T589E, N471D, N471D/T589D, and
 CC N471D/T589E. Also included are a glucan produced by the GFP mutant,
 CC an isolated polynucleotide which encodes P1 or P2, or its complementary
 CC polynucleotide, a ribonucleic acid sequence encoding the GFP mutant,
 CC an expression cassette comprising the polynucleotide operably linked to a
 CC promoter, a vector comprising the expression cassette, host cell
 CC introduced with the vector, a transgenic plant comprising the
 CC vector, a seed or tuber from the transgenic plant, a paper sizing and/or
 CC coating composition comprising a glucan produced in a plant transformed
 CC with a gene encoding the mutant GFP, wild type or, starch and starch where
 CC the thermoplastic molecule or their combinations or glucan and starch where
 CC the glucan is produced in the amyloplast and/or vacuole or a maize line
 CC deficient in starch biosynthesis, transformed with a gene encoding a
 CC glucosyltransferase B or D enzyme, wild-type or mutant and a paper
 CC comprising the glucan (paper sizing/coating agent). The vector is useful
 CC for producing a glucan in a plant. The method comprises transforming a
 CC plant cell with the vector, growing the plant cell under plant growing
 CC conditions to produce a regenerated plant and inducing expression of the
 CC polynucleotide for a time sufficient to produce the glucan in the
 CC regenerated plant, where the vector contains a transit sequence from
 CC ribulose biphosphate carboxylase small subunit, waxy, brittle-1 and
 CC chlorophyll AB binding protein to produce a transgenic plant, and glucan
 CC is produced in the amyloplast of potato or the vacuole of sugar beet.
 CC Glucans are useful as substitutes for and additions to modified starch
 CC and latexes in paper manufacture. Unlike prior art techniques, which
 CC require input materials that produce chemical effluents, paper
 CC manufacture utilizing the glucan produced by GFP, which utilizes
 CC biologically produced input materials, is more cost-effective and
 CC environmentally friendly. Moreover, glucans also exhibit thermoplastic
 CC properties and impart gloss to the paper during coating step.
 CC The present sequence represents a GFP mutant of the invention.
 CC Note: The present sequence is not shown in the specification but
 CC was created by the indexer using the GFPB sequence appearing as AAU98027
 CC and the information in claim 36.

XX SQ Sequence 1475 AA;
 SQ Query Match 100.0%; Score 119; DB 23; Length 1475;
 Best Local Similarity 100.0%; Pred. No. 5.4e-09;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ANDHSLILEAWSNDNTPYLHD 21
 Db 481 ANDHSLILEAWSNDNTPYLHD 501

RESULT 12
 AAU98039 ID AAU98039 standard; Protein; 1475 AA.
 XX AC AAU98039;
 XX

CC N471D/T569E. Also included are a glucan produced by the GTF mutant,
CC an isolated polynucleotide which encodes P1 or P2, or its complementary
CC polynucleotide, a ribonucleic acid sequence encoding the GTF mutant,
CC an expression cassette comprising the polynucleotide operably linked to a
CC promoter, a vector comprising the expression cassette, host cell
CC introduced with the vector, a transgenic plant comprising the
CC vector, a seed or tuber from the transgenic plant, a paper sizing and/or
CC coating composition comprising a glucan produced in a plant transformed
CC with a gene encoding the mutant GTF, wild type or, starch, a latex,
CC thermoplastic molecule or their combinations or glucan and starch where
CC the glucan is produced in the amyloplast and/or vacuole or a maize line
CC deficient in starch biosynthesis, transformed with a gene encoding a
CC glucosyltransferase B or D enzyme, wild-type or mutant and a paper
CC comprising the glucan (paper sizing/coating agent). The vector is useful
CC for producing a glucan in a plant. The method comprises transforming a
CC plant cell with the vector, growing the plant cell under plant growing
CC conditions to produce a regenerated plant and inducing expression of the
CC polynucleotide for a time sufficient to produce the glucan in the
CC regenerated plant, where the vector contains a transit sequence from
CC ribulose biphosphate carboxylase small subunit, waxy, brittle-1 and
CC chlorophyll A/B binding protein to produce a transgenic plant, and glucan
CC is produced in the amyloplast of potato or the vacuole of sugar beet.
CC Glucans are useful as substitutes for and additions to modified starch
CC and latexes in paper manufacture. Unlike prior art techniques, which
CC require input materials that produce chemical effluents, paper
CC manufacture utilizing the glucan produced by GTF, which utilizes
CC biologically produced input materials, is more cost-effective and
CC environmentally friendly. Moreover, glucans also exhibit thermoplastic
CC properties and impart gloss to the paper during coating step.
CC The present sequence represents a GTFB mutant of the invention.
CC Note: The present sequence is not shown in the specification but
CC was created by the indexer using the GTFB sequence appearing as AAU98027
CC and the information in claim 36.

CC Sequence 1475 AA;

Query Match 100.0%; Score 119; DB 23; Length 1475;
Best Local Similarity 100.0%; Pred. No. 5.4e-09;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ANDHLSILEAWSNDPTYLHD 21
|||
DB 481 ANDHLSILEAWSNDPTYLHD 501

RESULT 14
AAU79284
ID AAU79284 standard; Protein; 1476 AA.

AC AAU79284;
DT 13-AUG-2002 (first entry)

DE Streptococcus mutans monoclonal antibody-related protein #1.

KM Antibody; dental caries; water insoluble glucan synthetase;
KW anti-carries; glucosyl transferase-B; immunotherapy.

OS Streptococcus mutans.

PN JP2002114709-A.

PD 16-APR-2002.

PF 04-OCT-2000; 2000JP-0304889.

PR 04-OCT-2000; 2000JP-0304889.

XX (UNIT-) UNITV NIPPON.

DR WPI; 2002-448101/48.

PT Anti-carries agent composed of a monoclonal antibody against an

PT inhibitory enzyme against water insoluble glucan synthetase of glucosyl
PT transferase-B (GTF-B) of Streptococcus mutans -
XX
PS Claim 3; Page 13-16; 28pp; Japanese.

XX The invention relates to a monoclonal antibody against dental caries and
CC an anti-carries agent composed of a monoclonal antibody produced by
CC Streptococcus mutans, particularly mouse-hybridoma MHP126 (FERM P-17566)
CC or mouse-hybridoma MHP136 (FERM P-17567), against an enzyme having
CC inhibitive activity against water insoluble glucan synthetase of glucosyl
CC transferase-B. The monoclonal antibody specifically inhibits water
CC insoluble glucan synthetase of Streptococcus mutans produced glucosyl
CC transferase-B and is used in the immunotherapy of dental caries. This
CC sequence represents a Streptococcus mutans monoclonal antibody-related
CC protein.

XX Sequence 1476 AA;

Query Match 100.0%; Score 119; DB 23; Length 1476;
Best Local Similarity 100.0%; Pred. No. 5.4e-09;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ANDHLSILEAWSNDPTYLHD 21
|||
DB 481 ANDHLSILEAWSNDPTYLHD 501

RESULT 15
AAU98028
ID AAU98028 standard; Protein; 1375 AA.

AC AAU98028;

DT 27-AUG-2002 (first entry)

DE S. mutans glucosyltransferase GTFC.

KM Glucosyltransferase; GTFC; transgenic plant; paper sizing;
KW coating composition; glucan; starch; latex; thermoplastic molecule;
KW amyloplast; vacuole; paper manufacture.

OS Streptococcus mutans.

PN US2002031826-A1.

PD 14-MAR-2002.

PF 19-DEC-2000; 2000US-0740274.

PR 11-DEC-1998; 98US-0210361.

PR 07-JUN-1995; 95US-0478704.

PR 07-JUN-1995; 95US-0482711.

PR 07-JUN-1995; 95US-0485243.

PR 16-JAN-1998; 98US-0007999.

PR 16-JAN-1998; 98US-0008172.

PR 20-JAN-1998; 98US-0009620.

XX (NICH/) NICHOLS S E.

XX Nichols SE;

XX WPI; 2002-414332/44.

DR N-PSDB; ABR52939.

PT Glucosyltransferase B or D protein useful for producing a glucan useful
as substitutes for and additions to modified starch and latexes in
paper manufacture, comprises mutations in specific positions -
XX Disclosure; Page 30-33; 44pp; English.

CC The invention an isolated protein comprising a glucosyltransferase
CC (GTF) B polypeptide having changes at position 1448V, D457N,
CC D567Y, K1014I, D457N/D567I, D457N/D571K, D567Y/D571K,

CC D567T/D571K/K1014T, 1448V/D457N/D567T/D571K/K779Q/K1014T,
 CC Y169A/Y170A/Y171A, and K779Q or a GFP D polypeptide having
 CC changes at positions from T589D, T589E, N471D, N471D/T589D, and
 CC N471D/T589E. Also included are a glucan produced by the GFP mutant,
 CC an isolated polynucleotide which encodes P1 or P2, or its complementary
 CC polynucleotide, a ribonucleic acid sequence encoding the GFP mutant,
 CC an expression cassette comprising the polynucleotide operably linked to a
 CC promoter, a vector comprising the expression cassette, host cell
 CC introduced with the vector, a transgenic plant comprising the
 CC vector, a seed or tuber from the transgenic plant, a paper sizing and/or
 CC coating composition comprising a glucan produced in a plant transformed
 CC with a gene encoding the mutant GFP, wild type or, starch, a latex,
 CC thermoplastic molecule or their combinations or glucan and starch where
 CC the glucan is produced in the amyloplast and/or vacuole or a maize line
 CC deficient in starch biosynthesis, transformed with a gene encoding a
 CC glucosyltransferase B or D enzyme, wild-type or mutant and a paper
 CC comprising the glucan (paper sizing/coating agent). The vector is useful
 CC for producing a glucan in a plant. The method comprises transforming a
 CC plant cell with the vector, growing the plant cell under plant growing
 CC conditions to produce a regenerated plant and inducing expression of the
 CC polynucleotide for a time sufficient to produce the glucan in the
 CC regenerated plant, where the vector contains a transit sequence from
 CC ribulose biphosphate carboxylase small subunit, waxy, brittle-1 and
 CC chlorophyll AB binding protein to produce a transgenic plant, and glucan
 CC is produced in the amyloplast of potato or the vacuole of sugar beet.
 CC Glucans are useful as substitutes for and additions to modified starch
 CC and latexes in paper manufacture. Unlike prior art techniques, which
 CC require input materials that produce chemical effluents, paper
 CC manufacture utilizing the glucan produced by GFP, which utilizes
 CC biologically produced input materials, is more cost-effective and
 CC environmentally friendly. Moreover, glucans also exhibit thermoplastic
 CC properties and impart gloss to the paper during coating step.
 CC The present sequence represents GFPc.
 CC XX

SQ Sequence 1375 AA;

Query Match 92.4%; Score 110; DB 23; Length 1375;

Best Local Similarity 95.2%; Pred. No. 1.1e-07;

Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ANDHLSILEAWSNDPTPYLHD 21
 |||||
 Db 507 ANDHLSILEAWSNDPTPYLHD 527

Search completed: November 13, 2003, 09:38:25
 Job time : 39.5166 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 13, 2003, 09:45:40 ; Search time 27.4692 Seconds
(without alignments)
139.565 Million cell updates/sec

Title: US-09-290-049a-1
Perfect score: 119
Sequence: 1 ANDHLSILEAMSDNDTPYLHD 21

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 666188 seqs, 182559486 residues

Total number of hits satisfying chosen parameters: 666188

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Published Applications AA:*
1: /cgn2_6/prodata/2/pubppaa/US07_PUBCOMB.pep:*
2: /cgn2_6/prodata/2/pubppaa/PCM_NEW_PUB.pep:*
3: /cgn2_6/prodata/2/pubppaa/US06_NEW_PUB.pep:*
4: /cgn2_6/prodata/2/pubppaa/US06_PUBCOMB.pep:*
5: /cgn2_6/prodata/2/pubppaa/US07_NEW_PUB.pep:*
6: /cgn2_6/prodata/2/pubppaa/PCMUS_PUBCOMB.pep:*
7: /cgn2_6/prodata/2/pubppaa/US08_NEW_PUB.pep:*
8: /cgn2_6/prodata/2/pubppaa/US08_PUBCOMB.pep:*
9: /cgn2_6/prodata/2/pubppaa/US09_PUBCOMB.pep:*
10: /cgn2_6/prodata/2/pubppaa/US09_PUBCOMB.pep:*
11: /cgn2_6/prodata/2/pubppaa/US09_NEW_PUB.pep:*
12: /cgn2_6/prodata/2/pubppaa/US10_PUBCOMB.pep:*
13: /cgn2_6/prodata/2/pubppaa/US10_PUBCOMB.pep:*
14: /cgn2_6/prodata/2/pubppaa/US10_PUBCOMB.pep:*
15: /cgn2_6/prodata/2/pubppaa/US10_PUBCOMB.pep:*
16: /cgn2_6/prodata/2/pubppaa/US10_NEW_PUB.pep:*
17: /cgn2_6/prodata/2/pubppaa/US60_NEW_PUB.pep:*
18: /cgn2_6/prodata/2/pubppaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	119	100.0	1475	9	US-09-740-274-2
2	110	92.4	1375	9	US-09-740-274-4
3	75	63.0	545	10	US-09-995-749A-10
4	75	63.0	1430	9	US-09-740-274-6
5	71	59.7	522	10	US-09-995-749A-11
6	60	50.4	584	10	US-09-995-749A-12
7	51	42.9	302	12	US-10-009-823A-1
8	49	41.2	302	9	US-09-815-242-14004
9	49	41.2	535	10	US-09-995-749A-13
10	49	41.2	1781	10	US-09-995-749A-2
11	47.5	39.9	534	10	US-09-344-882-24
12	47.5	39.9	534	10	US-10-293-865-24
13	47	39.5	15	12	US-10-295-693-90
14	47	39.5	15	15	US-10-067-649-90
15	46.5	39.1	1604	9	US-09-888-615-73

16	46	38.7	199	10	US-09-738-626-4571	Sequence 4571, Ap
17	46	38.7	476	11	US-09-908-299-9	Sequence 34, Appl
18	46	38.7	539	12	US-10-230-026-34	Sequence 9, Appl
19	45.5	38.2	156	12	US-10-001-245-125	Sequence 125, App
20	45	37.8	268	15	US-10-156-761-12823	Sequence 12823, A
21	44.5	37.4	210	15	US-10-127-032-167	Sequence 167, App
22	44	37.0	216	15	US-10-156-761-9158	Sequence 1158, Ap
23	44	37.0	322	9	US-09-815-242-11634	Sequence 11634, A
24	44	37.0	331	9	US-09-815-242-10188	Sequence 10188, A
25	44	37.0	334	9	US-09-815-242-113981	Sequence 13981, A
26	44	37.0	749	15	US-10-211-962-96	Sequence 96, Appl
27	44	37.0	816	15	US-10-080-114A-7	Sequence 7, Appl
28	43.5	36.6	155	12	US-10-001-245-126	Sequence 126, App
29	43	36.1	41	11	US-09-764-891-3308	Sequence 3308, Ap
30	43	36.1	311	11	US-09-878-781-18	Sequence 18, Appl
31	43	36.1	358	11	US-09-095-478-9	Sequence 9, Appl
32	43	36.1	381	11	US-09-095-478-9	Sequence 7, Appl
33	43	36.1	715	15	US-10-156-761-7562	Sequence 7562, Ap
34	43	36.1	1267	15	US-10-059-585-56	Sequence 56, Appl
35	43	36.1	2466	12	US-10-177-980-12	Sequence 12, Appl
36	43	36.1	2485	9	US-09-802-669-46	Sequence 46, Appl
37	43	36.1	2654	15	US-10-227-610-2	Sequence 2, Appl
38	42.5	35.7	491	9	US-09-815-242-10190	Sequence 10190, A
39	42.5	35.7	491	9	US-09-815-242-13756	Sequence 13756, A
40	42	35.3	31	15	US-10-092-908-40	Sequence 40, Appl
41	42	35.3	41	15	US-10-092-908-41	Sequence 41, Appl
42	42	35.3	95	9	US-09-945-301-14	Sequence 14, Appl
43	42	35.3	298	11	US-09-988-067B-26	Sequence 26, Appl
44	42	35.3	348	10	US-09-738-626-6855	Sequence 6855, Ap
45	42	35.3	484	15	US-10-156-761-8339	Sequence 8339, Ap

ALIGNMENTS

RESULT 1
US-09-740-274-2
Sequence 2, Application US/09740274
Patent No. US20020031826A1
GENERAL INFORMATION:
APPLICANT: Nicholls, Scott E.
TITLE OF INVENTION: Glucan-containing Compositions and Paper
FILE REFERENCE: 0357CRD
CURRENT APPLICATION NUMBER: US/09/740,274
CURRENT FILING DATE: 2000-12-19
PRIOR APPLICATION NUMBER: 09/210,361
PRIOR FILING DATE: 1998-12-11
PRIOR APPLICATION NUMBER: 09/007,999
PRIOR FILING DATE: 1998-01-16
PRIOR APPLICATION NUMBER: 08/478,704
PRIOR FILING DATE: 1995-06-07
PRIOR APPLICATION NUMBER: 09/009,620
PRIOR FILING DATE: 1998-01-20
PRIOR APPLICATION NUMBER: 08/485,243
PRIOR FILING DATE: 1995-06-07
PRIOR APPLICATION NUMBER: 09/008,172
PRIOR FILING DATE: 1998-01-16
PRIOR APPLICATION NUMBER: 08/482,711
PRIOR FILING DATE: 1995-06-07
NUMBER OF SEQ ID NOS: 6
SOFTWARE: FaetsSQ for Windows Version 3.0
SEQ ID NO 2
LENGTH: 1475
TYPE: PRT
ORGANISM: Streptococcus mutans
US-09-740-274-2

Query Match 100.0%; Score 119; DB 9; Length 1475;
Best local Similarity 100.0%; Pred. No. 1.8e-09;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 ANDHLSILEAMSDNDTPYLHD 21

Db 481 ANDHLSTLEAWSNDPTPLHD 501

RESULT 2

US-09-740-274-4
; Sequence 4, Application US/09740274
; Patent No. US20020031826A1
; GENERAL INFORMATION:
; APPLICANT: Nichols, Scott E.
; TITLE OF INVENTION: Glucan-containing Compositions and Paper
; FILE REFERENCE: 0357CRD
; CURRENT APPLICATION NUMBER: US/09/740,274
; PRIOR FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: 09/210,361
; PRIOR FILING DATE: 1998-12-11
; PRIOR APPLICATION NUMBER: 09/007,999
; PRIOR FILING DATE: 1998-01-16
; PRIOR APPLICATION NUMBER: 08/478,704
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: 09/009,620
; PRIOR FILING DATE: 1998-01-20
; PRIOR APPLICATION NUMBER: 08/485,243
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: 09/008,172
; PRIOR FILING DATE: 1998-01-16
; PRIOR APPLICATION NUMBER: 08/482,711
; PRIOR FILING DATE: 1995-06-07
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 1375
; TYPE: PRT
; ORGANISM: streptococcus mutans
US-09-740-274-4

Query Match 92.4%; Score 110; DB 9; Length 1375;
Best Local Similarity 95.2%; Pred. No. 4e-08; 1; Indels 0; Gaps 0;
Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 ANDHLSTLEAWSNDPTPLHD 21
Db 507 ANDHLSTLEAWSNDPTPLHD 527

RESULT 3

US-09-995-749A-10
; Sequence 10, Application US/09995749A
; Patent No. US20020155568A1
; GENERAL INFORMATION:
; APPLICANT: VAN GEEL-SCHUTTEN, GERRITDINA HENDRIKA
; APPLICANT: DIJKHUIZEN, LUBBERT
; APPLICANT: RAHAOUI, HAKIM
; APPLICANT: LEER, ROBERT-JAN
; TITLE OF INVENTION: NOVEL GLUCOSYLTRANSFERASES
; FILE REFERENCE: B043388-CIP
; CURRENT APPLICATION NUMBER: US/09/995,749A
; PRIOR FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: 09/604,957
; PRIOR FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: EPO 00201871.1
; PRIOR FILING DATE: 2000-05-25
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
; LENGTH: 545
; TYPE: PRT
; ORGANISM: Streptococcus mutans
US-09-995-749A-10

Query Match 63.0%; Score 75; DB 10; Length 545;
Best Local Similarity 71.4%; Pred. No. 0.0033;
Matches 15; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 ANDHLSTLEAWSNDPTPLHD 21
Db 75 AINHLSTLEAWSNDPTQYKND 95

RESULT 4

US-09-740-274-6
; Sequence 6, Application US/09740274
; Patent No. US20020031826A1
; GENERAL INFORMATION:
; APPLICANT: Nichols, Scott E.
; TITLE OF INVENTION: Glucan-containing Compositions and Paper
; FILE REFERENCE: 0357CRD
; CURRENT APPLICATION NUMBER: US/09/740,274
; PRIOR FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: 09/210,361
; PRIOR FILING DATE: 1998-12-11
; PRIOR APPLICATION NUMBER: 09/007,999
; PRIOR FILING DATE: 1998-01-16
; PRIOR APPLICATION NUMBER: 08/478,704
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: 09/009,620
; PRIOR FILING DATE: 1998-01-20
; PRIOR APPLICATION NUMBER: 08/485,243
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: 09/008,172
; PRIOR FILING DATE: 1998-01-16
; PRIOR APPLICATION NUMBER: 08/482,711
; PRIOR FILING DATE: 1995-06-07
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 6
; LENGTH: 1430
; TYPE: PRT
; ORGANISM: streptococcus mutans
US-09-740-274-6

Query Match 63.0%; Score 75; DB 9; Length 1430;
Best Local Similarity 71.4%; Pred. No. 0.0097;
Matches 15; Conservative 1; Mismatches 5; Indels 0; Gaps 0;
QY 1 ANDHLSTLEAWSNDPTPLHD 21
Db 495 AINHLSTLEAWSNDPTQYKND 515

RESULT 5

US-09-995-749A-11
; Sequence 11, Application US/09995749A
; Patent No. US20020155568A1
; GENERAL INFORMATION:
; APPLICANT: VAN GEEL-SCHUTTEN, GERRITDINA HENDRIKA
; APPLICANT: DIJKHUIZEN, LUBBERT
; APPLICANT: RAHAOUI, HAKIM
; APPLICANT: LEER, ROBERT-JAN
; TITLE OF INVENTION: NOVEL GLUCOSYLTRANSFERASES
; FILE REFERENCE: B043388-CIP
; CURRENT APPLICATION NUMBER: US/09/995,749A
; PRIOR FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: 09/604,957
; PRIOR FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: EPO 00201871.1
; PRIOR FILING DATE: 2000-05-25
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 11
; LENGTH: 522
; TYPE: PRT
; ORGANISM: Leuconostoc mesenteroides
US-09-995-749A-11

Query Match 59.7%; Score 71; DB 10; Length 522;
Best Local Similarity 66.7%; Pred. No. 0.013;

Matches 14; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 1 ANDHSLTEAMSNDTPYLD 21
Db 75 ANQHLSTLEDSHNDPLXYTD 95

RESULT 6
US-09-995-749A-12

; Sequence 12, Application US/09995749A
; Patent No. US20020155568A1
; GENERAL INFORMATION:
; APPLICANT: VAN GEEL-SCHUTTEN, GERRITDINA HENDRIKA
; APPLICANT: DIJKHUITZEN, LUBBERT
; APPLICANT: RAHAOUI, HAKIM
; APPLICANT: LEER, ROBERT-JAN
; TITLE OF INVENTION: NOVEL GLUCOSYLTRANSFERASES
; FILE REFERENCE: B043388-CIP
; CURRENT APPLICATION NUMBER: US/09/995,749A
; CURRENT FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: 09/604,957
; PRIOR FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: EPO 00201871.1
; PRIOR FILING DATE: 2000-05-25
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 12
; LENGTH: 584
; TYPE: PRT
; ORGANISM: Leuconostoc mesenteroides
US-09-995-749A-12

Query Match 50.4%; Score 60; DB 10; Length 584;
Best Local Similarity 55.0%; Pred. No. 0.72;

Matches 11; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 1 ANDHSLTEAMSNDTPYLD 20
Db 75 ANKHLSTLEDSHNDPLXYTD 94

RESULT 7
US-10-009-823A-1

; Sequence 1, Application US/10009823A
; Publication No. US20030157120A1
; GENERAL INFORMATION:
; APPLICANT: Panaccio, Michael
; APPLICANT: Rosey, Everett Lee
; APPLICANT: Sinistraj, Meri
; APPLICANT: Haase, Detlef
; APPLICANT: Parsons, Jim
; APPLICANT: Ankenbauer, Robert G.
; TITLE OF INVENTION: LAMSONIA DERIVED GENE AND RELATED FLGE
; FILE REFERENCE: DAVI150.001APC
; CURRENT APPLICATION NUMBER: US/10/009,823A
; CURRENT FILING DATE: 2002-08-13
; PRIOR APPLICATION NUMBER: PCT/AU00/00437
; PRIOR FILING DATE: 2000-05-11
; PRIOR APPLICATION NUMBER: US 60/133,973
; PRIOR FILING DATE: 1999-05-13
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 502
; TYPE: PRT
; ORGANISM: Lamsonia intracellularis
US-10-009-823A-1

Query Match 42.9%; Score 51; DB 12; Length 502;
Best Local Similarity 47.1%; Pred. No. 14;
Matches 8; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 1 ANDHSLTEAMSNDTPYLD 17
Db 185 ANPYFALLESMKNGTNP 201

RESULT 8
US-09-815-242-14004

; Sequence 14004, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: BLITRA.01A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 1410
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14004
; LENGTH: 302
; TYPE: PRT
; ORGANISM: Salmonella typhi
US-09-815-242-14004

Query Match 41.2%; Score 49; DB 9; Length 302;
Best Local Similarity 50.0%; Pred. No. 17;
Matches 10; Conservative 1; Mismatches 9; Indels 0; Gaps 0;

QY 2 NDHSLTEAMSNDTPYLD 21
Db 267 NDRVLIVDEVLEASDIPYRD 266

RESULT 9
US-09-995-749A-13

; Sequence 13, Application US/09995749A
; Patent No. US20020155568A1
; GENERAL INFORMATION:
; APPLICANT: VAN GEEL-SCHUTTEN, GERRITDINA HENDRIKA
; APPLICANT: DIJKHUITZEN, LUBBERT
; APPLICANT: RAHAOUI, HAKIM
; APPLICANT: LEER, ROBERT-JAN
; TITLE OF INVENTION: NOVEL GLUCOSYLTRANSFERASES
; FILE REFERENCE: B043388-CIP
; CURRENT APPLICATION NUMBER: US/09/995,749A
; CURRENT FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: 09/604,957
; PRIOR FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: EPO 00201871.1
; PRIOR FILING DATE: 2000-05-25
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 13
LENGTH: 535
TYPE: PRT
ORGANISM: Lactobacillus reuteri
US-09-995-749A-13

Query Match 41.2%; Score 49; DB 10; Length 535;
Best Local Similarity 40.0%; Pred. No. 31;
Matches 8; Conservative 5; Mismatches 7; Indels 0; Gaps 0;

QY 1 ANDHSLTEAMSNDTPYLH 20
DB 74 SNKHINILEDWVHADPEYFN 93

RESULT 10
US-09-995-749A-2
Sequence 2, Application US/09995749A
Patent No. US2002015568A1
GENERAL INFORMATION:
APPLICANT: VAN GEL-SCHUTTEN, GERRITDINA, HENDRIKA
APPLICANT: DIJKHUIZEN, LUBBERT
APPLICANT: RAHAOUI, HAKIM
APPLICANT: LEER, ROBERT-JAN
TITLE OF INVENTION: NOVEL GLUCOSYLTRANSFERASES
FILE REFERENCE: BO43388-CIP
CURRENT APPLICATION NUMBER: US/09/995,749A
CURRENT FILING DATE: 2001-11-29
PRIOR FILING DATE: 2000-06-28
PRIOR APPLICATION NUMBER: EPO 00201871.1
PRIOR FILING DATE: 2000-05-25
NUMBER OF SEQ ID NOS: 19
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 2
LENGTH: 1781
TYPE: PRT
ORGANISM: Lactobacillus reuteri
US-09-995-749A-2

Query Match 41.2%; Score 49; DB 10; Length 1781;
Best Local Similarity 40.0%; Pred. No. 1,2e+02;
Matches 8; Conservative 5; Mismatches 7; Indels 0; Gaps 0;

QY 1 ANDHSLTEAMSNDTPYLH 20
DB 1053 SNKHINILEDWVHADPEYFN 1072

RESULT 11
US-09-344-882-24
Sequence 24, Application US/09344882
Patent No. US20020162137A1
GENERAL INFORMATION:
APPLICANT: Nikolau, Basil J
APPLICANT: Wurttele, Eve S
APPLICANT: Oliver, David J
APPLICANT: Behal, Robert
APPLICANT: Schnable, Patrick S
APPLICANT: Ke, Jinsuan
APPLICANT: Johnson, Jerry L
APPLICANT: Allred, Carolyn C
APPLICANT: Fatland, Beth
APPLICANT: Lutziger, Isabelle
TITLE OF INVENTION: Materials and Methods for the Alteration of Enzyme and
TITLE OF INVENTION: Acetyl CoA Levels in Plants
FILE REFERENCE: 201573
CURRENT APPLICATION NUMBER: US/09/344,882
CURRENT FILING DATE: 1999-06-25
PRIOR FILING DATE: 1998-06-26
NUMBER OF SEQ ID NOS: 38

SOFTWARE: PatentIn Ver. 2.2
SEQ ID NO 24
LENGTH: 534
TYPE: PRT
ORGANISM: Arabidopsis thaliana
US-09-344-882-24

Query Match 39.9%; Score 47.5; DB 10; Length 534;
Best Local Similarity 52.9%; Pred. No. 53;
Matches 9; Conservative 2; Mismatches 5; Indels 1; Gaps 1;

QY 2 NDHSLTEAMSNDTPY 18
DB 132 NDHSLTEW-DNGKPY 147

RESULT 12
US-10-293-865-24
Sequence 24, Application US/10293865
Publication No. US20030106090A1
GENERAL INFORMATION:
APPLICANT: Nikolau, Basil J
APPLICANT: Wurttele, Eve S
APPLICANT: Oliver, David J
APPLICANT: Behal, Robert
APPLICANT: Schnable, Patrick S
APPLICANT: Ke, Jinsuan
APPLICANT: Johnson, Jerry L
APPLICANT: Allred, Carolyn C
APPLICANT: Fatland, Beth
APPLICANT: Lutziger, Isabelle
TITLE OF INVENTION: Materials and Methods for the Alteration of Enzyme and
TITLE OF INVENTION: Acetyl CoA Levels in Plants
FILE REFERENCE: 217113
CURRENT APPLICATION NUMBER: US/10/293,865
CURRENT FILING DATE: 2002-11-13
PRIOR FILING DATE: 1999-06-25
PRIOR APPLICATION NUMBER: US 09/344,882
PRIOR FILING DATE: 1998-06-26
NUMBER OF SEQ ID NOS: 38
SOFTWARE: PatentIn Ver. 3.1
SEQ ID NO 24
LENGTH: 534
TYPE: PRT
ORGANISM: Arabidopsis thaliana
US-10-293-865-24

Query Match 39.9%; Score 47.5; DB 15; Length 534;
Best Local Similarity 52.9%; Pred. No. 53;
Matches 9; Conservative 2; Mismatches 5; Indels 1; Gaps 1;

QY 2 NDHSLTEAMSNDTPY 18
DB 132 NDHSLTEW-DNGKPY 147

RESULT 13
US-10-295-693-90
Sequence 90, Application US/10295693
Publication No. US20030198976A1
GENERAL INFORMATION:
APPLICANT: Bristol-Myers Squibb Company
TITLE OF INVENTION: NOVEL HUMAN G-PROTEIN COUPLED RECEPTOR, HGRPRWY14, RELATED TO THE
TITLE OF INVENTION: GPCR, GPCR3
FILE REFERENCE: D0118 CIP
CURRENT APPLICATION NUMBER: US/10/295,693
CURRENT FILING DATE: 2002-11-14
PRIOR FILING DATE: 2001-02-05
PRIOR APPLICATION NUMBER: US 60/266,525
PRIOR FILING DATE: 2002-02-05

PRIOR APPLICATION NUMBER: US 60/329,897
PRIOR FILING DATE: 2001-10-16
NUMBER OF SEQ ID NOS: 95
SOFTWARE: PatentIn version 3.1
SEQ ID NO 90
LENGTH: 15
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Synthesized Random Peptide.
US-10-295-693-90

Query Match 39.5%; Score 47; DB 12; Length 15;
Best Local Similarity 61.5%; Pred. No. 1.2; 2; Indels 0; Gaps 0;
Matches 8; Conservative 3; Mismatches 0;

QY 8 LEAMSDNDTPYLH 20
Db 3 LEAMDLSDTPHLV 15

RESULT 14
US-10-067-649-90
Sequence 90, Application US/10067649
Publication No. US20030100057A1
GENERAL INFORMATION:
APPLICANT: Bristol-Myers Squibb Company
TITLE OF INVENTION: A NOVEL HUMAN G-PROTEIN COUPLED RECEPTOR, HGPRTMY14, RELATED TO 1
FILE REFERENCE: D0118 NP
CURRENT APPLICATION NUMBER: US/10/067,649
CURRENT FILING DATE: 2002-02-05
PRIOR APPLICATION NUMBER: US 60/266,525
PRIOR FILING DATE: 2001-02-05
PRIOR APPLICATION NUMBER: US 60/329,897
PRIOR FILING DATE: 2001-10-16
NUMBER OF SEQ ID NOS: 92
SOFTWARE: PatentIn version 3.0
SEQ ID NO 90
LENGTH: 15
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Synthesized Random Peptide.
US-10-067-649-90

Query March 39.5%; Score 47; DB 15; Length 15;
Best Local Similarity 61.5%; Pred. No. 1.2; 2; Indels 0; Gaps 0;
Matches 8; Conservative 3; Mismatches 0;

QY 8 LEAMSDNDTPYLH 20
Db 3 LEAMDLSDTPHLV 15

RESULT 15
US-09-888-615-73
Sequence 73, Application US/09888615
Patent No. US20020064856A1
GENERAL INFORMATION:
APPLICANT: PLOWMAN, GREGORY
APPLICANT: WHITE, DAVID
APPLICANT: CAENEPEEL, SEAN
APPLICANT: CHARYDOZAK, GLEN
APPLICANT: MANNING, GERARD
APPLICANT: SUDARSANAM, SUCHA
TITLE OF INVENTION: NOVEL PROTEASES
FILE REFERENCE: 038602/1214
CURRENT APPLICATION NUMBER: US/09/888,615
CURRENT FILING DATE: 2001-06-26
PRIOR APPLICATION NUMBER: 60/214,047
PRIOR FILING DATE: 2000-06-26
NUMBER OF SEQ ID NOS: 150

SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 73
LENGTH: 1604
TYPE: PRT
ORGANISM: Homo sapiens
US-09-888-615-73

Query Match 39.1%; Score 46.5; DB 9; Length 1604;
Best Local Similarity 47.4%; Pred. No. 2.6e+02;
Matches 9; Conservative 3; Mismatches 4; Indels 3; Gaps 1;

QY 5 ISILEAMSDN---DTPYLH 20
Db 294 VALLEWVKDNRRTDDIPELH 312

Search completed: November 13, 2003, 10:28:59
Job time : 28.4692 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 13, 2003, 09:31:40 ; Search time 7.56398 seconds
(without alignments)
130.561 Million cell updates/sec

Title: US-09-290-049a-1
Perfect score: 119
Sequence: 1 ANDHLSTLEAWSNDNTPYLHD 21

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues
Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	119	100.0	1476	1	GTPB_STRMU P08987 streptococc
2	110	92.4	1455	1	GTF2_STRMU P13470 streptococc
3	108	90.8	1592	1	GTF2_STRDO P27470 streptococc
4	108	90.8	1597	1	GTF1_STRDO P11001 streptococc
5	78	65.5	1365	1	GTF5_STRDO P29336 streptococc
6	75	63.0	1462	1	GTFD_STRMU P49331 streptococc
7	57	47.9	330	1	GTPC_LETME P01558 leishmania
8	49	41.2	302	1	YBHK_SALTI P58586 salmonella
9	49	41.2	302	1	YBHK_SALTY P58586 salmonella
10	49	41.2	336	1	UL16_EBV P03221 Epstein-Barr
11	47.5	39.9	330	1	G3PC_TRYBB P10097 trypanosoma
12	47	39.5	337	1	G3P_COCP0 P51113 coccidioid
13	46	38.7	275	1	SC35_YEAST P51271 saccharomyc
14	46	38.7	335	1	G3P_CHUPN P09270 chlamydia p
15	46	38.7	476	1	PCHA_PSEAE P01508 pseudomonas
16	45	37.8	236	1	GAMT_HUMAN Q14353 homo sapien
17	45	37.8	385	1	GUNF_FUSOX P46239 fusarium ox
18	44.5	37.4	470	1	AMP2_STRCO P06039 streptomyc
19	44	37.0	234	1	G3P_MEIGA P06972 escherichia
20	44	37.0	330	1	G3PI_ECOLI P24165 salmonella
21	44	37.0	330	1	G3PI_SALTY P00356 gallus gall
22	44	37.0	332	1	G3P_CHICK P05479 columba liv
23	44	37.0	332	1	G3P_COOLI P08735 zea mays (m
24	44	37.0	337	1	G3PD_MAIZE P00904 zea mays (m
25	44	37.0	337	1	G3PD_HORVU P26511 hordeum vul
26	44	37.0	337	1	G3P_CRYPA P15089 cryptonectr
27	44	37.0	337	1	G3P_CRYPA P15089 cryptonectr
28	44	37.0	363	1	G3PI_JACOR P49036 jaculus ori
29	44	37.0	816	1	SUS2_MAIZE P41540 oryctolagus
30	43.5	36.6	115	1	TNKL_RABIT P06054 mesocricetu
31	43.5	36.6	130	1	TNKL_MEATU P41539 mus musculu
32	43.5	36.6	130	1	TNKL_MOUSE P06767 rattus norv
33	43.5	36.6	130	1	TNKL_MOUSE P06767 rattus norv

34	43.5	36.6	491	1	G6PD_ERMCH P37986 erwania chr
35	43	36.1	174	1	PADC_VIRCH O9KRX2 vibrio chol
36	43	36.1	235	1	GAMT_RAT P10868 rattus norv
37	43	36.1	236	1	GAMT_MOUSE O35969 mus muscula
38	43	36.1	320	1	G3P_BOVIN P10096 bos taurus
39	43	36.1	322	1	G3P_SHEEP O28554 ovis aries
40	43	36.1	332	1	G3P_PIG P00355 sus scrofa
41	43	36.1	332	1	G3P_PIG P00355 sus scrofa
42	43	36.1	334	1	G3P2_HUMAN P04406 homo sapien
43	43	36.1	337	1	G3P_AJECA O9HFX1 ajeilomyces
44	43	36.1	337	1	G3P_MYCGB P47543 mycoplasma
45	43	36.1	424	1	IYD_RAT P12007 rattus norv

ALIGNMENTS

RESULT 1
GTFB_STRMU STANDARD; PRT; 1476 AA.
ID GTFB_STRMU
AC P08987; O69381; O69384; O69387; O69390; O69396;
DT 01-NOV-1988 (Rel. 09, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE (Glucosyltransferase-1 precursor (EC 2.4.1.5) (GTF-1) (Dextranucrase)
GN GTFB OR SMU.1004.
OS Streptococcus mutans.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OX NCBI_TaxId=1309;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=GS-5;
RX MEDLINE=87308013; PubMed=3040685;
RA Shiroza T., Ueda S., Kuramitsu H.K.;
RT "Sequence analysis of the gtfB gene from Streptococcus mutans.";
RL J. Bacteriol. 169:4263-4270(1987).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=MT4239 / Serotype C, MT4245 / Serotype E, MT4251 / Serotype F,
RC MT4467 / Serotype B, and MT8148 / Serotype C;
RX MEDLINE=98231643; PubMed=9570124;
RA Fujiwara T., Terao Y., Hoshino T., Kawabata S., Ooshima T., Sobue S.,
RA Kimura S., Hamada S.;
RT "Molecular analyses of glucosyltransferase genes among strains of
RT Streptococcus mutans";
RL FEMS Microbiol. Lett. 161:331-336(1998).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=DA159 / ATCC 700610 / Serotype C;
RX MEDLINE=22295063; PubMed=12397186;
RA Ajdic D., McShan W.M., McLaughlin R.E., Savic G., Chang J.,
RA Carson M.B., Primeaux C., Tian R., Kenton S., Jia H., Jin S., Qian Y.,
RA Li S., Zhu H., Najjar F., Lai H., White J., Roe B.A., Ferretti J.J.;
RT "Genome sequence of Streptococcus mutans DA159, a cariogenic dental
RT pathogen";
RL Proc. Natl. Acad. Sci. U.S.A. 99:14434-14439(2002).
RN [4]
RP PROC. NATL. ACADEMY OF SCIENCES
RX MEDLINE=12397186; PubMed=12397186;
RA "TO PLAY A KEY ROLE IN THE DEVELOPMENT OF THE DENTAL PLAQUE BECAUSE
OF THEIR ABILITY TO ADHERE TO SMOOTH SURFACES AND MEDIATE THE
AGGREGATION OF BACTERIAL CELLS AND FOOD DEBRIS."
CC -1- CATALYTIC ACTIVITY: Sucrose + (1,6)-alpha-D-glucosyl (N) = D-
CC -1- fructose + (1,6)-alpha-D-glucosyl (N+1).
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- DISEASE: DENTAL CARIES.
CC -1- MISCELLANEOUS: GTF-1 SYNTHESIZES WATER-INSOLUBLE GLUCANS (ALPHA
CC 1,3-LINKED GLUCOSE AND SOME 1,6 LINKAGS). GTF-SI SYNTHESIZES BOTH
CC WATER-SOLUBLE GLUCANS (ALPHA 1,6-GLUCOSE).
CC FORMS OF GLUCANS.
CC -1- SIMILARITY: TO OTHER GLUCOSYLTRANSFERASES AND SOME TO A GLUCAN-
CC BINDING PROTEIN FROM S. MUTANS.
CC -1- SIMILARITY: Contains 10 cell wall binding repeats.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; M17361; AAA89588.1; -
 DR EMBL; D88651; BAA26101.1; -
 DR EMBL; D88654; BAA26105.1; -
 DR EMBL; D88657; BAA26109.1; -
 DR EMBL; D88660; BAA26113.1; -
 DR EMBL; D88677; BAA26119.1; -
 DR EMBL; AE014940; AAN58705.1; -
 DR InterPro; IPR002479; CW binding.
 DR InterPro; IPR003318; Glyco_hydro_70.
 DR Pfam; PF01473; CW binding_I; 1.
 DR Pfam; PF02324; Glyco_hydro_70; 1.
 KW Transferase; Glycosyltransferase; Signal; Repeat; Dental caries;
 KW Complete proteome.
 FT SIGNAL 1 34
 FT CHAIN 35 1476
 FT DOMAIN 35 1051
 FT DOMAIN 1097 1476
 FT REPEAT 1097 1130
 FT DOMAIN 1161 1470
 FT REPEAT 1161 1210
 FT REPEAT 1225 1275
 FT REPEAT 1290 1340
 FT REPEAT 1355 1405
 FT REPEAT 1420 1470
 FT REPEAT 62 62
 FT VARIANT 65 65
 FT VARIANT 68 68
 FT VARIANT 78 78
 FT VARIANT 86 86
 FT VARIANT 89 89
 FT VARIANT 158 168
 FT VARIANT 276 276
 FT VARIANT 399 399
 FT VARIANT 474 474
 FT VARIANT 512 512
 FT VARIANT 519 519
 FT VARIANT 701 701
 FT VARIANT 708 708
 FT VARIANT 938 938
 FT VARIANT 952 952
 FT VARIANT 963 964
 FT VARIANT 968 970
 FT VARIANT 1086 1086
 FT VARIANT 1158 1158
 FT VARIANT 1163 1163
 FT VARIANT 1168 1168
 FT VARIANT 1182 1182
 FT VARIANT 1234 1234
 FT VARIANT 1263 1263
 FT VARIANT 1264 1264
 FT VARIANT 1264 1264
 FT VARIANT 1272 1272
 FT VARIANT 1329 1329
 FT VARIANT 1394 1394
 FT VARIANT 1402 1402

POTENTIAL.
 GLUCOSYLTRANSFERASE-1.
 CATALYTIC (APPROXIMATE).
 GLUCAN-BINDING (APPROXIMATE).
 A REPEAT.
 5 X TANDEM REPEATS.
 1.
 2.
 3.
 4.
 5.
 S -> T (IN STRAIN MT4239).
 T -> I (IN STRAIN GS-5).
 V -> A (IN STRAINS GS-5, MT4245, MT4251,
 MT4467 AND MT8148).
 Q -> P (IN STRAIN MT4251).
 I -> S (IN STRAINS GS-5, MT4245, MT4251,
 MT4467 AND MT8148).
 S -> F (IN STRAIN MT4251).
 K -> N (IN STRAIN MT4251).
 S -> D (IN STRAINS GS-5, MT4467 AND
 MT8148).
 N -> R (IN STRAIN MT4239).
 I -> T (IN STRAIN MT4239).
 K -> R (IN STRAIN MT8148).
 F -> Y (IN STRAIN MT8148).
 T -> I (IN STRAIN MT8148).
 A -> V (IN STRAIN MT8148).
 F -> L (IN STRAIN MT8148).
 FGGPVE -> YGTPVA (IN STRAINS GS-5, MT4239
 AND MT4467).
 SV -> NT (IN STRAINS GS-5, MT4239 AND
 MT4467).
 ADS -> VDG (IN STRAINS GS-5, MT4239 AND
 MT4467).
 A -> T (IN STRAIN MT4239).
 S -> N (IN STRAIN MT4239).
 H -> Y (IN STRAIN MT4251).
 E -> K (IN STRAIN MT8148).
 Y -> C (IN STRAIN MT8148).
 A -> P (IN STRAIN MT4239).
 R -> P (IN STRAIN MT8148).
 R -> H (IN STRAINS GS-5 AND MT4467).
 Y -> H (IN STRAINS GS-5, MT4239, MT4467
 AND MT8148).
 S -> G (IN STRAINS GS-5, MT4239, MT4467
 AND MT8148).
 H -> Y (IN STRAINS GS-5 AND MT4467).
 Y -> H (IN STRAINS GS-5, MT4239, MT4467
 AND MT8148).
 S -> G (IN STRAINS GS-5, MT4239, MT4467

FT AND MT8148).
 FT VARIANT 1459 1459 Y -> H (IN STRAIN MT4467).
 FT CONFLICT 570 570 R -> A (IN REF. 1).
 FT CONFLICT 800 817 ADDVAVAASTAPSTDGK -> LKKFAURLARPHOQMA
 (IN REF. 1).
 FT CONFLICT 1310 1310 H -> L (IN REF. 1).
 FT SEQUENCE 1476 AA; 165846 MW; 9C6E09F731B4CBCE CRC64;
 SQ
 Query Match 100.0%; Score 119; DB 1; Length 1476;
 Best Local Similarity 100.0%; Pred. No. 1,2e-10;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 ANDHSLTEAWSDNDTPYLHD 21
 Db 481 ANDHSLTEAWSDNDTPYLHD 501
 RESULT 2
 ID GTFPC STRNM STANDARD; PRT; 1455 AA.
 AC P13470; O69382; O69385; O69388; O69391; O69397; P05427;
 DT 01-NOV-1988 (Rel. 09, Created)
 DT 26-FEB-2003 (Rel. 41, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Glucosyltransferase-ST precursor (EC 2.4.1.5) (GTF-ST)
 DE (dextranucrase) (sucrose 6-glucosyltransferase).
 GN GTFPC OR SMU.1005.
 OS Streptococcus mutans.
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 OX NCBI_TaxID=1309;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=GS-5;
 RX MEDLINE=89137980; PubMed=2976010;
 RA Ueda S., Shiroza T., Kuramitsu H.K.;
 RT "Sequence analysis of the gtfC gene from Streptococcus mutans GS-5.";
 RL Gene 69:101-109(1988).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MT4239 / Serotype C, MT4245 / Serotype E, MT4251 / Serotype F,
 RX MT4467 / Serotype E and MT8148 / Serotype C;
 RX MEDLINE=98231643; PubMed=9770124;
 RA Fujiwara T., Terao Y., Hoshino T., Kawabata S., Ooshima T., Sobue S.,
 RA Kimura S., Hamada S.;
 RT "Molecular analyses of glucosyltransferase genes among strains of
 RT Streptococcus mutans.";
 RL FEMS Microbiol. Lett. 161:331-336(1998).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=UAI59 / ATCC 700610 / Serotype C;
 RX MEDLINE=22295063; PubMed=12397186;
 RA Ajdic D., McShan W.M., McLaughlin R.E., Savic G., Chang J., Qian Y.,
 RA Carson M.B., Primeaux C., Tian R., Kenton S., Jia H., Lin S., Ferretti J.J.;
 RA Li S., Zhu H., Najjar F., Lai H., White J., Roe B.A., Ferretti J.J.;
 RT "Genome sequence of Streptococcus mutans UAI59, a cariogenic dental
 RT pathogen.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:14434-14439(2002).
 RN [4]
 RP SEQUENCE OF 1-349 FROM N.A.
 RC STRAIN=GS-5;
 RX MEDLINE=87308013; PubMed=3040685;
 RA Shiroza T., Ueda S., Kuramitsu H.K.;
 RT "Sequence analysis of the gtfB gene from Streptococcus mutans.";
 RL J. Bacteriol. 169:4263-4270(1987).
 CC -1- FUNCTION: PRODUCTION OF EXTRACELLULAR GLUCANS, THAT ARE THOUGHT
 CC TO PLAY A KEY ROLE IN THE DEVELOPMENT OF THE DENTAL PLAQUE BECAUSE
 CC OF THEIR ABILITY TO ADHERE TO SMOOTH SURFACES AND MEDATE THE
 CC AGGREGATION OF BACTERIAL CELLS AND FOOD DEBRIS.
 CC -1- CATALYTIC ACTIVITY: Sucrose + ((1,6)-alpha-D-glucosyl)(N) = D-
 CC fructose + ((1,6)-alpha-D-glucosyl)(N+1).
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- DISEASE: DENTAL CARIES.

CC -1- MISCELLANEOUS: GTF-1 SYNTHESIZES WATER-INSOLUBLE GLUCANS (ALPHA
CC 1,3-LINKED GLUCOSE AND SOME 1,6 LINKAGES), GTF-S SYNTHESIZES
CC WATER-SOLUBLE GLUCANS (ALPHA 1,6-GLUCOSE), GTF-SI SYNTHESIZES BOTH
CC FORMS OF GLUCANS.
CC -1- SIMILARITY: TO OTHER GLUCOSYLTRANSFERASES AND SOME TO A GLUCAN-
CC BINDING PROTEIN FROM S. MUTANS.
CC -1- SIMILARITY: Contains 5 cell wall binding repeats.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M2054; AAA88592.1; -
DR EMBL; D88652; BAA26102.1; -
DR EMBL; D88655; BAA26106.1; -
DR EMBL; D88658; BAA26110.1; -
DR EMBL; D88661; BAA26114.1; -
DR EMBL; D89978; BAA26120.1; -
DR EMBL; AE014940; AAN58706.1; -
DR EMBL; M17361; AAA88589.1; -
DR InterPro; IPR002479; CW_binding.
DR InterPro; IPR003318; Glyco_hydro_70.
DR Pfam; PF01473; CW_binding_1; 1.
DR Pfam; PF02324; Glyco_hydro_70; 1.
DR Transferase; Glycosyltransferase; Signal; Repeat; Dental caries;
KM Complete proteome.
KM SIGNAL 1 34
FT CHAIN 1 1455
FT DOMAIN 35 1050
FT DOMAIN 1126 1455
FT DOMAIN 1126 1455
FT REPEAT 1126 1159
FT REPEAT 1169 1200
FT REPEAT 1227 1238
FT REPEAT 1253 1303
FT REPEAT 1318 1330
FT REPEAT 21 21
FT VARIANT 81 81
FT VARIANT 106 106
FT VARIANT 116 116
FT VARIANT 126 126
FT VARIANT 150 151
FT VARIANT 256 256
FT VARIANT 425 425
FT VARIANT 519 519
FT VARIANT 538 538
FT VARIANT 545 545
FT VARIANT 597 597
FT VARIANT 600 600
FT VARIANT 601 601
FT VARIANT 614 614
FT VARIANT 727 727
FT VARIANT 734 734
FT VARIANT 964 964
FT VARIANT 1113 1113
FT VARIANT 1118 1118
FT VARIANT 1204 1204
FT VARIANT 1208 1208
FT VARIANT 1292 1294
FT VARIANT 1305 1369
FT VARIANT 1326 1326
FT VARIANT 1331 1331
FT VARIANT 1377 1377
GLUCOSYLTRANSFERASE-SI.
CATALYTIC (APPROXIMATE).
GLUCAN-BINDING (APPROXIMATE).
2-4 A, 1 C AND 1 AC REPEATS.
A REPEAT.
A REPEAT.
C REPEAT.
AC REPEAT.
A REPEAT (INCOMPLETE).
V -> I (IN STRAIN GS-5).
P -> L (IN STRAIN MT4239).
D -> V (IN STRAIN GS-5).
S -> A (IN STRAINS GS-5 AND MT4467).
A -> T (IN STRAIN GS-5).
SR -> PK (IN STRAINS GS-5, MT4239 AND
MT4467).
A -> V (IN STRAINS GS-5 AND MT4467).
R -> N (IN STRAIN MT4251).
Y -> D (IN STRAINS MT4245 AND MT4251).
R -> K (IN STRAINS MT4245 AND MT4251).
Y -> F (IN STRAINS MT4245 AND MT4251).
N -> D (IN STRAINS MT4245, MT4251, MT4467
AND MT8148).
R -> K (IN STRAINS MT4245, MT4251, MT4467
AND MT8148).
A -> T (IN STRAIN GS-5).
M -> T (IN STRAIN GS-5).
T -> I (IN STRAIN MT8148).
A -> V (IN STRAIN MT8148).
L -> F (IN STRAIN MT4239).
N -> Y (IN STRAIN MT4239).
A -> T (IN STRAIN MT4239).
I -> V (IN STRAINS GS-5, MT4239, MT4467
AND MT8148).
V -> I (IN STRAIN MT8148).
DGH -> NGY (IN STRAINS GS-5, MT4467 AND
MT8148).
MISSING (IN STRAIN MT4245).
I -> V (IN STRAIN MT8148).
T -> A (IN STRAINS GS-5, MT4239, MT4467
AND MT8148).
R -> K (IN STRAIN MT8148).

FT VARIANT 1398 1398 V -> I (IN STRAIN MT8148).
FT VARIANT 1424 1424 D -> N (IN STRAIN MT4239).
FT VARIANT 1439 1439 V -> I (IN STRAINS MT4239 AND MT8148).
FT VARIANT 1444 1444 S -> P (IN STRAIN MT8148).
FT CONFLICT 1337 1455 QLYFKSNGVQAKGELITERKRIKXYDPSNGEVNRYVR
TSSGMYYFGNDGALIGWVVEGRVYDEGVYVYASHD
ORNMVDYRDGRDGSAAVRFRSRNPFDFRFE ->
HASTISWVERLRRESLSQSVKVSNMILLPEKKFYVM
(IN REF. 1).
FT FT
SQ SEQUENCE 1455 AA; 162965 MW; 3CB455A99A4FEC86 CRC64;
FT FT
Query Match 92.4%; Score 110; DB 1; Length 1455;
Best Local Similarity 95.2%; Pred. No. 3.3e-09;
Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 ANDHLSILEAWSDNDTPYLHD 21
DB 507 ANDHLSILEAWSDNDTPYLHD 527
RESULT 3
ID GTF2_STRDO STANDARD; PRT; 1592 AA.
AC P27470;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Glucosyltransferase-I precursor (EC 2.4.1.5) (GTF-1) (Dextranucrase)
DE (Sucrose 6-glucosyltransferase).
OS Streptococcus downei (Streptococcus sobrinus).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NOBI_taxid=1317;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=6715 / Serotype G;
RX MEDLINE=91123227; PubMed=1704006;
RA Abo H., Matsumura T., Kodama T., Ohta H., Fukui K., Kato K.,
KA Kagawa H.;
RT "peptide sequences for sucrose splitting and glucan binding within
RT Streptococcus sobrinus glucosyltransferase (water-insoluble glucan
RT synthetase).";
RL Bacteriol. 173:989-996 (1991).
CC -1- FUNCTION: PRODUCTION OF EXTRACELLULAR GLUCANS, THAT ARE THOUGHT
CC TO PLAY A KEY ROLE IN THE DEVELOPMENT OF THE DENTAL PLAQUE BECAUSE
CC OF THEIR ABILITY TO ADHERE TO SMOOTH SURFACES AND MEDIANE THE
CC AGGREGATION OF BACTERIAL CELLS AND FOOD DEBRIS.
CC -1- CATALYTIC ACTIVITY: Sucrose + (1,6'-alpha-D-glucosyl) (N) = D-
CC fructose + (1,6)-alpha-D-glucosyl (N+1).
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- DISEASE: DENTAL CARIES.
CC -1- MISCELLANEOUS: GTF-1 SYNTHESIZES WATER-INSOLUBLE GLUCANS (ALPHA
CC 1,3-LINKED GLUCOSE AND SOME 1,6 LINKAGES), GTF-S SYNTHESIZES
CC WATER-SOLUBLE GLUCANS (ALPHA 1,6-GLUCOSE), GTF-SI SYNTHESIZES BOTH
CC FORMS OF GLUCANS.
CC -1- SIMILARITY: TO OTHER GLUCOSYLTRANSFERASES AND SOME TO A GLUCAN-
CC BINDING PROTEIN FROM S. MUTANS.
CC -1- SIMILARITY: Contains 16 cell wall binding repeats.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; D90213; BAA14241.1; -
DR InterPro; IPR002479; CW_binding.
DR InterPro; IPR003318; Glyco_hydro_70.
DR Pfam; PF01473; CW_binding_1; 13.
DR Pfam; PF02324; Glyco_hydro_70; 1.
DR Transferase; Glycosyltransferase; Signal; Repeat; Dental caries.
KM

```

FT SIGNAL 1 38 POTENTIAL.
FT CHAIN 39 1592 GLUCOSYLTRANSFERASE-1.
FT DOMAIN 39 1044 CATALYTIC (APPROXIMATE).
FT DOMAIN 1093 1592 GLUCAN-BINDING (APPROXIMATE).
FT DOMAIN 1093 1592 6.5 X TANDEM REPEATS.
FT REPEAT 1093 1142 1.
FT REPEAT 1158 1207 2.
FT REPEAT 1222 1272 3.
FT REPEAT 1287 1337 4.
FT REPEAT 1402 1451 5.
FT REPEAT 1514 1563 6.
FT REPEAT 1577 1592 7 (INCOMPLETE).
SQ SEQUENCE 1592 AA; 176167 MW; BCOA6D079351ECF CRC64;

Query Match 90.8%; Score 108; DB 1; Length 1592;
Best Local Similarity 85.7%; Pred. No. 7.6e-09;
Matches 18; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

OY 1 ANDHLSTEAMSDNDTPYLHD 21
DB 477 ANNHVSIVEAMSDNDTPYLHD 497

RESULT 4
GTF1_STRDO STANDARD; PRT: 1597 AA.
ID GTF1_STRDO
AC p11001;
DT 01-JUL-1989 (Rel. 11, Created)
DT 01-JUL-1989 (Rel. 11, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Glucosyltransferase-I precursor (EC 2.4.1.5) (GTF-I) (Dextranucrase)
DE (Sucrose 6-glucosyltransferase).
GN GTFI.
OS Streptococcus downei (Streptococcus sobrinus).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1317;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MFE28;
RX MEDLINE=87308014; PubMed=3040686;
RA Ferretti J.J., Gilpin M.L., Russell R.R.B.;
RT "Nucleotide sequence of a glucosyltransferase gene from Streptococcus
RT sobrinus MFE28."
RL J. Bacteriol. 169:4271-4278(1987).
CC - FUNCTION: PRODUCTION OF EXTRACELLULAR GLUCANS, THAT ARE THOUGHT
CC TO PLAY A KEY ROLE IN THE DEVELOPMENT OF THE DENTAL PLAQUE BECAUSE
CC OF THEIR ABILITY TO ADHERE TO SMOOTH SURFACES AND MEDIATE THE
CC AGGREGATION OF BACTERIAL CELLS AND FOOD DEBRIS.
CC - CATALYTIC ACTIVITY: Sucrose + (1,6)-alpha-D-glucosyl (N) = D-
CC fructose + (1,6)-alpha-D-glucosyl (N+1).
CC - SUBCELLULAR LOCATION: Secreted.
CC - DISEASE: DENTAL CARIES.
CC - MISCELLANEOUS: GTF-I SYNTHESIZES WATER-INSOLUBLE GLUCANS (ALPHA
CC 1,3-LINKED GLUCOSE AND SOME 1,6 LINKAGES), GTF-S SYNTHESIZES
CC WATER-SOLUBLE GLUCANS (ALPHA 1,6-GLUCOSE). GTF-SI SYNTHESIZES BOTH
CC FORMS OF GLUCANS.
CC - SIMILARITY: TO OTHER GLUCOSYLTRANSFERASES AND SOME TO A GLUCAN-
CC BINDING PROTEIN FROM S.MUTANS.
CC - SIMILARITY: Contains 19 cell wall binding repeats.
CC - This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M17391; AAC63063.1; -
CC InterPro; IPR002479; CW_binding.
CC InterPro; IPR003318; Glyco_hydro_70.
CC Pfam; PF01473; CW_binding_1; 16.

```

```

DR Pfam; PF02324; Glyco_hydro_70; 1.
KM Transferase; Glucosyltransferase; Signal; Repeat; Dental caries.
FT SIGNAL 1 38 POTENTIAL.
FT CHAIN 39 1597 GLUCOSYLTRANSFERASE-1.
FT DOMAIN 39 1050 CATALYTIC (APPROXIMATE).
FT DOMAIN 1099 1597 GLUCAN-BINDING (APPROXIMATE).
FT DOMAIN 1099 1597 1.25 A, 2 B, AND 5 AC REPEATS.
FT REPEAT 1099 1132 A REPEAT.
FT REPEAT 1163 1213 AC REPEAT.
FT REPEAT 1227 1277 AC REPEAT.
FT REPEAT 1292 1342 AC REPEAT.
FT REPEAT 1352 1399 B REPEAT.
FT REPEAT 1406 1455 AC REPEAT.
FT REPEAT 1465 1512 B REPEAT.
FT REPEAT 1519 1568 AC REPEAT.
FT REPEAT 1582 1597 A REPEAT (INCOMPLETE).
SQ SEQUENCE 1597 AA; 177080 MW; B9E86A200868798E CRC64;

Query Match 90.8%; Score 108; DB 1; Length 1597;
Best Local Similarity 85.7%; Pred. No. 7.6e-09;
Matches 18; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

OY 1 ANDHLSTEAMSDNDTPYLHD 21
DB 483 ANNHVSIVEAMSDNDTPYLHD 503

RESULT 5
GTF5_STRDO STANDARD; PRT: 1365 AA.
ID GTF5_STRDO
AC p29336;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Glucosyltransferase-S precursor (EC 2.4.1.5) (GTF-S) (Dextranucrase)
DE (Sucrose 6-glucosyltransferase).
GN GTF5.
OS Streptococcus downei (Streptococcus sobrinus).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1317;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MFE28;
RX MEDLINE=90316665; PubMed=2142479;
RA Gilmore K.S., Russell R.R., Ferretti J.J.;
RT "Analysis of the Streptococcus downei gtf5 gene, which specifies a
RT glucosyltransferase that synthesizes soluble glucans."
RL Infect. Immun. 58:2452-2458(1990).
CC - FUNCTION: PRODUCTION OF EXTRACELLULAR GLUCANS, THAT ARE THOUGHT
CC TO PLAY A KEY ROLE IN THE DEVELOPMENT OF THE DENTAL PLAQUE BECAUSE
CC OF THEIR ABILITY TO ADHERE TO SMOOTH SURFACES AND MEDIATE THE
CC AGGREGATION OF BACTERIAL CELLS AND FOOD DEBRIS.
CC - CATALYTIC ACTIVITY: Sucrose + (1,6)-alpha-D-glucosyl (N) = D-
CC fructose + (1,6)-alpha-D-glucosyl (N+1).
CC - ENZYME REGULATION: GLUCAN SYNTHESIS BY GTF-S IS INDEPENDENT OF
CC PRIMER GLUCAN UNLIKE GTF-I.
CC - DISEASE: DENTAL CARIES.
CC - MISCELLANEOUS: GTF-S SYNTHESIZES WATER-SOLUBLE GLUCANS (ALPHA
CC 1,6-GLUCOSE).
CC - SIMILARITY: TO OTHER GLUCOSYLTRANSFERASES AND SOME TO A GLUCAN-
CC BINDING PROTEIN FROM S.MUTANS.
CC - SIMILARITY: Contains 10 cell wall binding repeats.
CC - This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M30943; AAA26898.1; -

```

DR InterPro: IPR002479; CW binding.
 DR InterPro: IPR003318; Glyco_hydro_70.
 DR Pfam: PF02324; Glyco_hydro_70; 1.
 DR Pfam: PF02324; Glycosyltransferase; Signal; Repeat; Dental caries.
 FT TRANSFERASE; Glycosyltransferase; Signal; Repeat; Dental caries.
 FT SIGNAL 1 36 OR 37 (POTENTIAL).
 FT CHAIN 37 1365 GLUCOSYLTRANSFERASE-S.
 FT REPEAT 157 177 CELL WALL BINDING 1.
 FT REPEAT 178 197 CELL WALL BINDING 2.
 FT DOMAIN 198 1061 CATALYTIC (APPROXIMATE).
 FT REPEAT 1062 1082 CELL WALL BINDING 3.
 FT REPEAT 1083 1102 CELL WALL BINDING 4.
 FT REPEAT 1150 1169 CELL WALL BINDING 5.
 FT REPEAT 1170 1190 CELL WALL BINDING 6.
 FT REPEAT 1225 1243 CELL WALL BINDING 7.
 FT REPEAT 1289 1308 CELL WALL BINDING 8.
 FT REPEAT 1309 1328 CELL WALL BINDING 9.
 FT REPEAT 1331 1352 CELL WALL BINDING 10.
 SQ SEQUENCE 1365 AA; 151590 MW; 16729685A2E8C476 CRC64;

Query Match 65.5%; Score 78; DB 1; Length 1365;
 Best Local Similarity 71.4%; Pred. NO. 0.00037;
 Matches 15; Conservatively 1; Mismatches 5; Indels 0; Gaps 0;

OY 1 ANDHLSTLEAMSNDPEYLD 21
 Db 467 AIDHLSTLEAMSNDNDYVD 487

RESULT 6
 GTFD STRMU STANDARD; PRT: 1462 AA.
 AC P49331; O69383; O69386; O69389; O69392; O69398;
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 28-FEB-2003 (Rel. 41, last sequence update)
 DT 15-SEP-2003 (Rel. 42, last annotation update)
 DE Glucosyltransferase-S precursor (EC 2.4.1.5) (GTF-S) (Dextranucrase)
 DE (Sucrose 6-glucosyltransferase).
 GN GTFD OR SMU.910.
 OS Streptococcus mutans.
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 OX NCBI_TaxId=1309;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=GS-5;
 RX MEDLINE=91100958; PubMed=2148600;
 RA Honda O., Kato C., Kuramitsu H.K.;
 RT "Nucleotide sequence of the Streptococcus mutans gtfD gene encoding the glucosyltransferase-S enzyme";
 RL J. Gen. Microbiol. 136:2099-2105(1990).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MT4239 / Serotype C, MT4245 / Serotype E, MT4251 / Serotype F,
 RC MT4467 / Serotype E, and MT8148 / Serotype C;
 RX MEDLINE=98231643; PubMed=9570124;
 RA Fujiwara T., Terao Y., Hoshino T., Kawabata S., Oshima T., Sobue S.,
 RA Kimura S., Hamada S.;
 RT "Molecular analyses of glucosyltransferase genes among strains of Streptococcus mutans";
 RL FEMS Microbiol. Lett. 161:331-336(1998).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=UA159 / ATCC 700610 / Serotype C;
 RX MEDLINE=22295063; PubMed=12397186;
 RA Ajdic D., McShan W.M., McLaughlin R.E., Savic G., Chang J., Qian Y.,
 RA Carson M.B., Primeaux C., Tian R., Kenton S., Jia H., Ish S., Qian Y.,
 RA Li S., Zhu H., Najjar F., Lai H., White J., Roe B.A., Ferretti J.J.;
 RT "Genome sequence of Streptococcus mutans UA159, a cariogenic dental pathogen";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:14434-14439(2002).
 CC -1- FUNCTION: PRODUCTION OF EXTRACELLULAR GLUCANS, THAT ARE THOUGHT TO PLAY A KEY ROLE IN THE DEVELOPMENT OF THE DENTAL PLAQUE BECAUSE

CC OF THEIR ABILITY TO ADHERE TO SMOOTH SURFACES AND MEDIATE THE
 CC AGGREGATION OF BACTERIAL CELLS AND FOOD DEBRIS.
 CC -1- CATALYTIC ACTIVITY: Sucrose + {(1,6)-alpha-D-glucosyl}(N) = D-
 CC fructose + {(1,6)-alpha-D-glucosyl}(N+1).
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- DISEASE: DENTAL CARIES.
 CC -1- MISCELLANEOUS: GTF-I SYNTHESIZES WATER-INSOLUBLE GLUCANS (ALPHA
 CC 1,3-LINKED GLUCOSE AND SOME 1,6 LINKAGES), GTF-S SYNTHESIZES
 CC WATER-SOLUBLE GLUCANS (ALPHA 1,6-GLUCOSE). GTF-SI SYNTHESIZES BOTH
 CC FORMS OF GLUCANS.
 CC -1- SIMILARITY: TO OTHER GLUCOSYLTRANSFERASES AND SOME TO A GLUCAN-
 CC BINDING PROTEIN FROM S.MUTANS.
 CC -1- SIMILARITY: Contains 6 cell wall binding repeats.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: M29286; AAA26895.1; -
 CC EMBL: D88653; BAA26103.1; -
 CC EMBL: D88656; BAA26107.1; -
 CC EMBL: D88659; BAA26111.1; -
 CC EMBL: D88662; BAA26115.1; -
 CC EMBL: D89979; BAA26121.1; -
 CC EMBL: A8014932; AAN58619.1; -
 DR InterPro: IPR002479; CW-binding.
 DR InterPro: IPR003318; Glyco_hydro_70.
 DR Pfam: PF02324; Glyco_hydro_70; 1.
 DR Pfam: PF02324; Glycosyltransferase; Signal; Repeat; Dental caries;
 KW Complete proteome.
 FT SIGNAL 1 1462
 FT CHAIN 1232 1423
 FT DOMAIN 1232 1295
 FT REPEAT 1296 1359
 FT REPEAT 1360 1423
 FT REPEAT 10 10
 FT VARIANT 19 19
 FT VARIANT 58 58
 FT VARIANT 68 68
 FT VARIANT 81 81
 FT VARIANT 113 113
 FT VARIANT 122 122
 FT VARIANT 132 132
 FT VARIANT 135 135
 FT VARIANT 137 137
 FT VARIANT 202 202
 FT VARIANT 255 255
 FT VARIANT 275 275
 FT VARIANT 288 288
 FT VARIANT 301 301
 FT VARIANT 313 313
 FT VARIANT 317 317
 FT VARIANT 328 328
 FT VARIANT 350 350
 FT VARIANT 628 633
 FT VARIANT 688 688
 FT VARIANT 726 732
 FT VARIANT 726 730
 FT VARIANT 762 762

POTENTIAL.
 GLUCOSYLTRANSFERASE-S.
 3 X 63 AA APPROXIMATE TANDEM REPEATS.
 1.
 2.
 3.
 Y -> H (IN STRAINS GS-5, MT4239, MT4245,
 MT4251, MT4467 AND MT8148)
 I -> V (IN STRAINS GS-5, MT4239, MT4245,
 MT4251, MT4467 AND MT8148).
 K -> E (IN STRAIN MT4467).
 A -> S (IN STRAINS MT4239 AND MT4245).
 A -> T (IN STRAINS MT4251 AND MT8148).
 I -> V (IN STRAINS MT4239, MT4245 AND
 MT8148).
 A -> S (IN STRAINS GS-5 AND MT4467).
 A -> T (IN STRAIN MT4245).
 A -> V (IN STRAINS GS-5, MT4239, MT4245,
 MT4251, MT4467 AND MT8148).
 V -> L (IN STRAIN MT4239).
 D -> N (IN STRAIN MT8148).
 E -> D (IN STRAINS MT4239, MT4245 AND
 MT4251).
 D -> N (IN STRAINS MT4239, MT4245 AND
 MT4251).
 Q -> H (IN STRAIN MT4245).
 D -> N (IN STRAINS MT4239 AND MT4251).
 E -> K (IN STRAIN MT4239).
 V -> F (IN STRAIN MT4239).
 F -> L (IN STRAINS MT4239, MT4251 AND
 MT4467).
 KKRTQ -> EKEETL (IN STRAIN MT4251).
 A -> S (IN STRAIN MT4239).
 TDQSEA -> ADKGNS (IN STRAIN MT4251).
 TDQGS -> ADKGN (IN STRAINS MT4239 AND
 MT4245).
 T -> A (IN STRAINS GS-5, MT4239, MT4245,
 MT4251, MT4467 AND MT8148).


```

FT VARIANT 964 964 D -> Y (IN STRAIN MT4251).
FT VARIANT 1019 1019 E -> K (IN STRAINS MT4245 AND MT4251).
FT VARIANT 1059 1060 LG -> R (IN STRAIN MT4251).
FT VARIANT 1060 1060 KM GLYCOLYSIS; Oxidoreductase; NAD.
FT VARIANT 1080 1080 G -> R (IN STRAIN MT4245).
FT VARIANT 1142 1142 H -> Q (IN STRAIN GS-5).
FT VARIANT 1142 1142 S -> N (IN STRAIN MT4239).
FT VARIANT 1198 1198 Y -> C (IN STRAINS MT4251 AND MT4467).
FT VARIANT 1220 1220 F -> L (IN STRAIN MT4467).
FT VARIANT 1280 1280 Q -> P (IN STRAIN MT4245).
FT VARIANT 1282 1282 K -> T (IN STRAIN MT4245).
FT VARIANT 1290 1290 N -> D (IN STRAIN MT4245).
FT VARIANT 1311 1311 G -> D (IN STRAINS GS-5 AND MT4467).
FT VARIANT 1403 1403 G -> R (IN STRAIN GS-5).
FT VARIANT 1425 1425 R -> K (IN STRAIN MT4467).
FT VARIANT 1449 1449 RYDKNSGMVYMKVTLANGRIIGDIRMGIRY -> VY
FT CONFLICT 1428 1462 R (IN REF. 1).
SQ SEQUENCE 1462 AA; 163387 MW; CE4A279C4D708645 CRC64;

```

```

Query Match 63.0%; Score 75; DB 1; Length 1462;
Best Local Similarity 71.4%; Pred. No. 0.0012;
Matches 15; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

```

```

Oy 1 ANDHSTLEAMSNDTPYLH 21
Db 495 AINHSTLEAMSNDTPQYK 515

```

```

RESULT 7
G3PC_LEIME STANDARD; PRT; 330 AA.
ID G3PC_LEIME
AC Q01558;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Glyceroldehyde 3-phosphate dehydrogenase, cytosolic (EC 1.2.1.12)
DE (GAPDH).
DE Leishmania mexicana.
OS Eukaryota; Eulendozoa; Kinetoplastida; Trypanosomatidae; Leishmania.
OC NCB1_Taxid=5665;
OX NCB1_Taxid=5665;
RN SEQUENCE FROM N.A.
RP STRAIN=ssp. Mexicana;
RX MEDLINE=93063042; PubMed=1435864;
RA Hannaert V., Blaauw M., Kohl L., Allert S., Oppendoes F.R.,
RA Michels P.A.M.;
RT "Molecular analysis of the cytosolic and glycosomal glyceroldehyde-3-
RT phosphate dehydrogenase in Leishmania mexicana.";
RT Mol. Biochem. Parasitol. 55:115-126(1992).
CC -1- CATALYTIC ACTIVITY: D-glyceroldehyde 3-phosphate + phosphate +
CC NAD(+) = 3-phospho-D-glyceroyl phosphate + NADH.
CC -1- PATHWAY: Second phase of glycolysis; first step.
CC -1- SUBUNIT: Homotetramer.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: Belongs to the glyceroldehyde 3-phosphate
CC dehydrogenase family.

```

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

```

CC EMBL; X65220; CAA46323.1; -.
CC PIR; B48445; B48445.
CC HSSP; P06977; IGAD.
CC InterPro; IPR000173; GAP_dhhydrogenase.
CC InterPro; IPR006424; GAPDH-I.
CC Pfam; PF00044; gpdh; 1.
CC Pfam; PF02800; gpdh_C; 1.

```

```

DR PRINTS; PR00078; G3PDHRCNASE.
DR TIGRfams; TIGR01534; GAPDH-I; 1.
DR PROSITE; PS00071; GAPDH; 1.
KM GLYCOLYSIS; Oxidoreductase; NAD.
FT INIT MET 0 BY SIMILARITY.
FT BINDING 148 148 GLYCERALDEHYDE 3-PHOSPHATE.
FT ACT SITE 175 175 ACTIVATES THIOL GROUP DURING CATALYSIS.
SQ SEQUENCE 330 AA; 35511 MW; ED4B6D85A207F1E CRC64;

```

```

Query Match 47.3%; Score 57; DB 1; Length 330;
Best Local Similarity 52.6%; Pred. No. 0.14;
Matches 10; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

```

```

Oy 2 NDHSTLEAMSNDTPYLH 20
Db 300 NDHVKVXVSWYDNETGYSH 318

```

```

RESULT 8
YBHK_SALTI STANDARD; PRT; 302 AA.
ID YBHK_SALTI
AC P58587;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Hypothetical UPF0052 protein ybhK.
DE YBHK OR S10835 OR T2086.
OS Salmonella typhi.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Salmonella.
OX NCB1_Taxid=601;
OX [1]
RN SEQUENCE FROM N.A.
RP STRAIN=CT18;
RX MEDLINE=21534947; PubMed=11677608;
RA Parkhill J., Dougan G., James K.D., Thomson N.R., Pickard D., Wain J.,
RA Churcher C., Mungall K.L., Bentley S.D., Holden M.T.G., Sebatina M.,
RA Baker S., Basham D., Brooks K., Chillingworth T., Connor P.,
RA Cronin A., Davis P., Davies R.M., Dowd L., White N., Farrar J.,
RA Felwell T., Hamlin N., Haque A., Hien T.T., Holroyd S., Jagsels K.,
RA Krogh A., Larsen T.S., Leachter S., Moule S., O'Gaora P., Parry C.,
RA Quail M., Rutherford K., Simmonds M., Skelton J., Stevens K.,
RA Whitehead S., Barrett B.G.;
RT "Complete genome sequence of a multiple drug resistant Salmonella
RT enterica serovar Typhi CT18."
RT Nature 413:848-852(2001).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=Ty2 / ATCC 700931;
RX MEDLINE=22531367; PubMed=12644504;
RA Deng W., Liou S.-R., Plunkett G. III, Mayhew G.F., Rose D.J.,
RA Burdand V., Kodoyianni V., Schwartz D.C., Blatner F.R.;
RT "Comparative genomics of Salmonella enterica serovar Typhi strains Ty2
RT and CT18."
RT J. Bacteriol. 185:2330-2337(2003).
CC -1- SIMILARITY: Belongs to the UPF0052 family.

```

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

```

CC EMBL; AL627268; CAD05249.1; -.
CC DR EMBL; AE016841; AAO69704.1; -.
CC DR InterPro; IPR02882; UPF0052.
CC DR Pfam; PF01933; UPF0052; 1.
CC KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 302 AA; 32656 MW; 2218471211E185C3 CRC64;

```

```

Query Match 41.2%; Score 49; DB 1; Length 302;

```

Best Local Similarity 50.0%; Pred. No. 2.4;
Matches 10; Conservative 1; Mismatches 9; Indels 0; Gaps 0;

QY 2 NDHLSILEAMSDNDPYLEHD 21
DB 267 NDRLVIOEVLASDIPYRHD 286

RESULT 9

YBHK_SALTY STANDARD; PRT; 302 AA.

AC 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Hypothetical UPF0052 protein ybhK.
GN YBHK OR STM0801.
OS Salmonella typhimurium.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Salmonella.
OX NCBI_TaxID=602;

RP SEQUENCE FROM N.A.
RC STRAIN=LT2 / SGC1412 / ATCC 700720;
RX MEDLINE=21534948; PubMed=11677609;
RA McClelland M., Sanderson K.R., Spieth J., Clifton S.W., Latreille P.,
RA Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D.,
RA Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,
RA Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,
RA Waterston R., Wilson R.K.;
RT "Complete genome sequence of Salmonella enterica serovar Typhimurium
LT2.";
RL Nature 413:852-856(2001).

CC - SIMILARITY: Belongs to the UPF0052 family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL: AE008733; AAL19738.1; -
DR StyGene; SG32227; ybhK.
DR InterPro; IPR002882; UPF0052.
DR Pfam; PF01933; UPF0052; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 302 AA; 32626 MW; 38A84708A1B185C3 CRC64;

Query Match 41.2%; Score 49; DB 1; Length 302;
Best Local Similarity 50.0%; Pred. No. 2.4;
Matches 10; Conservative 1; Mismatches 9; Indels 0; Gaps 0;

QY 2 NDHLSILEAMSDNDPYLEHD 21
DB 267 NDRLVIOEVLASDIPYRHD 286

RESULT 10
UL16_EBV STANDARD; PRT; 336 AA.

AC 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE BGLF2 protein.
GN BGLF2.
OS Epstein-Barr virus (strain B95-8) (Human herpesvirus 4).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Gammaherpesvirinae; Lymphocryptovirus.
OX NCBI_TaxID=10377;
RN [1]

RP SEQUENCE FROM N.A.
RX MEDLINE=84270667; PubMed=6087149;
RA Baer R., Bankier A.T., Biggin M.D., Deininger P.L., Farrell P.J.,
RA Gibson T.J., Hatfull G., Hudson G.S., Satchwell S.C., Seglin C.,
RA Tuffnell P.S., Barrett B.G.;
RT "DNA sequence and expression of the B95-8 Epstein-Barr virus genome.";
RL Nature 310:207-211(1984).

RP SEQUENCE FROM N.A.
RX MEDLINE=92113548; PubMed=1662696;
RA Chen M.R., Hsu T.Y., Lin S.W., Chen J.Y., Yang C.S.;
RT "Cloning and characterization of cDNA clones corresponding to
RT transcripts from the BamHI G region of the Epstein-Barr virus genome
RT and expression of BGLF2.";
RL J. Gen. Virol. 72:3047-3055(1991).

CC - SIMILARITY: BELONGS TO FAMILY THAT GROUPS TOGETHER HSV-1 UL16,
CC HSV-6 ORF11R, EBV-1 46, HCMV UL94, EBV BDLF2, AND VZV 44.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL: M60514; AAA45871.1; -
DR EMBL: V01555; CAA24831.1; -
DR EMBL: S77133; AAB21113.1; -
DR PIR; C43044; Q0B840.
DR InterPro; IPR004286; UL16_UL94.
DR Pfam; PF03044; UL16_UL94; 1.
KW Late protein.
SQ SEQUENCE 336 AA; 36888 MW; 840937A416D5584C CRC64;

Query Match 41.2%; Score 49; DB 1; Length 336;
Best Local Similarity 41.2%; Pred. No. 2.8;
Matches 7; Conservative 6; Mismatches 4; Indels 0; Gaps 0;

QY 1 ANDHLSILEAMSDNDP 17
DB 216 AGAHVILKGMTDSDP 232

RESULT 11
G3PC_TRYB STANDARD; PRT; 330 AA.
AC P10097;
DT 01-MAR-1989 (Rel. 10, Created)
DT 01-MAY-1992 (Rel. 22, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Glyceraldehyde 3-phosphate dehydrogenase, cytosolic (EC 1.2.1.12)
DE (GAPDH).

OS Trypanosoma brucei brucei.
OC Eukaryota; Euklenozoa; Kinetoplastida; Trypanosomatidae; Trypanosoma.
OX NCBI_TaxID=5702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=427;
RX MEDLINE=91249838; PubMed=2040303;
RA Michels P.A.M., Marchand M., Kohl L., Allert S., Wierenga R.K.,
RA Operdoes F.R.;
RT "The cytosolic and glycosomal isoenzymes of glyceraldehyde-3-phosphate
RT dehydrogenase in Trypanosoma brucei have a distant evolutionary
RT relationship.";
RL Eur. J. Biochem. 198:421-428(1991).

RP PRELIMINARY SEQUENCE OF 1-85.
RX MEDLINE=87161817; PubMed=3830153;
RA Misset O., van Beeumen J., Lambey A.M., van der Meer R.,
RA Operdoes F.R.;
RT "Glyceraldehyde-phosphate dehydrogenase from Trypanosoma brucei.
RT Comparison of the glycosomal and cytosolic isoenzymes.";

```

RL  Eur. J. Biochem. 162:501-507(1987).
CC  -1- CATALYTIC ACTIVITY: D-glyceralddehyde 3-phosphate + phosphate +
CC  NAD(+) = 3-phospho-D-glyceroyl phosphate + NADH.
CC  -1- PATHWAY: Second phase of glycolysis, first step.
CC  -1- SUBUNIT: Homotetramer.
CC  -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC  -1- SIMILARITY: Belongs to the glyceraldehyde 3-phosphate
CC  dehydrogenase family.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
DR  EMBL; X53472; CAA37568.1; -.
DR  PIR; S16091; DEUTGC.
DR  HSSP; P06977; 1GMD.
DR  InterPro; IPR000173; GAP_dhhydrogenase.
DR  InterPro; IPR006424; GAPDH-I.
DR  Pfam; PF00044; gpdh.1.
DR  Pfam; PF02800; gpdh.C; 1.
DR  PRINTS; PR00078; G3PDHGRNASE.
DR  TIGRFS; TIGR01534; GAPDH-I; 1.
DR  PROSITE; PS00071; GAPDH; 1.
KW  Glycolysis; Oxidoreductase; NAD.
FT  INIT_MET 0
FT  BINDING 149 149 GLYCERALDEHYDE 3-PHOSPHATE.
FT  ACT_SITE 176 176 ACTIVATES THIOL GROUP DURING CATALYSIS.
FT  VARIANT 72 72 R -> I.
FT  CONFLICT 56 60 Q -> T (IN REF. 2).
FT  CONFLICT 60 60 Q -> K (IN REF. 2).
SQ  SEQUENCE 330 AA; 35503 MW; 7E1F0AFDA091E7C5 CRC64;

Query Match 39.9%; Score 47.5; DB 1; Length 330;
Best Local Similarity 43.5%; Pred. No. 4.7;
Matches 10; Conservative 4; Mismatches 6; Indels 3; Gaps 1;

QY 2 NDHSLILEAWSDNDTPY---LHD 21
DB 301 NDHFKLVSWYDNEGYSNKVDH 323

RESULT 12
G3P_COCPO STANDARD; PRT; 337 AA.
AC Q8J1H3;
DT 15-SEP-2003 (Rel. 42, Created)
DT 15-SEP-2003 (Rel. 42, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Glyceraldehyde 3-phosphate dehydrogenase (EC 1.2.1.12) (GAPDH).
GN GPD.
OS Coccidioides posadasii.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Oxygenales; mitosporic Oxygenales; Coccidioides.
OX NCBI_TaxID=199306;
RN [1]
RP SEQUENCE FROM N.A.
RA HUNG C.-Y., YU J.-J., COLE G.T.;
RT "Isolation and evaluation of expression of the glyceraldehyde-3-
RT phosphate dehydrogenase gene of Coccidioides immitis."
RT Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
RL
CC -1- CATALYTIC ACTIVITY: D-glyceraldehyde 3-phosphate + phosphate +
CC NAD(+) = 3-phospho-D-glyceroyl phosphate + NADH.
CC -1- PATHWAY: Second phase of glycolysis, first step.
CC -1- SUBUNIT: Homotetramer (By similarity).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: Belongs to the glyceraldehyde 3-phosphate
CC dehydrogenase family.
CC -----

```

```

CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
DR  EMBL; AF288134; AAN76496.1; -.
DR  InterPro; IPR006424; GAPDH-I.
DR  InterPro; IPR000173; GAP_dhhydrogenase.
DR  Pfam; PF00044; gpdh.1.
DR  Pfam; PF02800; gpdh.C; 1.
DR  PRINTS; PR00078; G3PDHGRNASE.
DR  TIGRFS; TIGR01534; GAPDH-I; 1.
DR  PROSITE; PS00071; GAPDH; 1.
KW  Glycolysis; Oxidoreductase; NAD.
FT  BINDING 151 151
FT  ACT_SITE 178 178 ACTIVATES THIOL GROUP DURING CATALYSIS.
SQ  SEQUENCE 337 AA; 36400 MW; 62EA834CA536DAF2 CRC64;

Query Match 39.5%; Score 47; DB 1; Length 337;
Best Local Similarity 52.9%; Pred. No. 5.7;
Matches 9; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 2 NDHSLILEAWSDNDTPY 18
DB 303 NDHFKLVSWYDNEGY 319

RESULT 13
SC35_YEAST STANDARD; PRT; 275 AA.
ID SC35_YEAST
AC P53271;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE SEC35 protein.
GN SEC35 OR YGR120C OR G6324.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLIN=5288; /F11679;
RX MEDLIN=97197982; PubMed=9046098;
RA van Dyck L., Tettelein H., Purnelle B., Goffeau A.;
RT "An 18.3 kb DNA fragment from yeast chromosome VII carries four
RT unknown open reading frames, the gene for an Aen synthase, remnants
RT of Ty and three tRNA genes."
RL Yeast 13:171-176(1997).
RN [2]
RP FUNCTION, AND SUBCELLULAR LOCATION.
RX MEDLIN=98270888; PubMed=9606204;
RA Vanheenen S.M., Cao X., Japashin V.V., Barlowe C., Waters M.G.;
RT "Sec35p, a novel peripheral membrane protein, is required for ER to
RT Golgi vesicle docking."
RL U. Cell Biol. 141:1107-1119(1998).
RN [3]
RP SUBUNIT.
RX MEDLIN=21563418; PubMed=11703943;
RA Whyte J.R., Munro S.;
RT "The Sec34/35 Golgi transport complex is related to the exocyst,
RT defining a family of complexes involved in multiple steps of membrane
RT traffic."
RL Dev. Cell 1:527-537(2001).
CC -1- FUNCTION: Required for ER to Golgi vesicle docking. Not essential
CC for viability.
CC -1- SUBUNIT: Component of the Sec34/Sec35 complex which consists of
CC eight different proteins.
CC -1- SUBCELLULAR LOCATION: Membrane-associated.
CC -----

```

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL; X83099; CAA58155.1; -
 CC EMBL; 272905; CAA97130.1; -
 CC PIR; S55978; S55978.
 CC SGD; S0003352; SEC35.
 CC GO; GO:0017119; C:Golgi transport complex; IPI.
 CC GO; GO:0000300; C:periplasmic membrane protein of membrane fra. .; IDA.
 CC GO; GO:0005625; C:soluble fraction; IDA.
 CC GO; GO:0005515; F:protein binding activity; IPI.
 CC GO; GO:0006888; P:ER to Golgi transport; IGI.
 CC GO; GO:0000301; P:retrograde vesicle recycling within Golgi. .; IMP.
 CC Transport; Protein transport; Golgi stack; Membrane.
 CC SEQUENCE 275 AA; 31799 MW; AA102D086FF3AD7 CRC64;

Query Match 38.7%; Score 46; DB 1; Length 275;
 Best Local Similarity 50.0%; Pred. No. 6.5;
 Matches 7; Conservative 6; Mismatches 1; Indels 0; Gaps 0;

QY 7 ILEAMSDNTPYLH 20
 DB 217 ILESCANSNPYH 230

RESULT 14
 G3P_CHLPPN STANDARD; PRT; 335 AA.
 ID G3P_CHLPPN
 AC Q92770; Q9Q0H7;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Glyceraldehyde 3-phosphate dehydrogenase (EC 1.2.1.12) (GAPDH).
 GN GAP OR GAPA OR CPN0624 OR CP0123.
 OS Chlamydia pneumoniae (Chlamydia pneumoniae).
 OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia.
 NCBI_TaxID=83558;
 RX SEQUENCE FROM N.A.
 RC STRAIN=CWL029;
 RA Kaiman S., Mitchell W., Marathe R., Lammel C., Fan J., Hyman R.W.,
 RA Olinger L., Grimwood J., Davis R.W., Stephens R.S.;
 RA "Comparative genomes of Chlamydia pneumoniae and C. trachomatis.";
 RA Nat. Genet. 21:385-389(1999).
 RL [2]
 RN SEQUENCE FROM N.A.
 RP STRAIN=AR39;
 RX MEDLINE=20150255; PubMed=10684935;
 RA Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.F.,
 RA White O., Hickey E.K., Peterson J., Tetreback T., Berry K., Bass S.,
 RA Linher K., Weidman J., Kouri H., Craven B., Bowman C., Dodson R.,
 RA Gwin M., Nelson W., DeBoy R., Kolonay J., McClarty G., Salzberg S.L.,
 RA Eisen J., Fraser C.M.;
 RA "Genome sequences of Chlamydia trachomatis Mopn and Chlamydia
 RT pneumoniae AR39.";
 RL Nucleic Acids Res. 28:1397-1406(2000).
 RL [3]
 RN SEQUENCE FROM N.A.
 RP STRAIN=J138;
 RX MEDLINE=20330349; PubMed=10871362;
 RA Shirai M., Hirakawa H., Kimoto M., Tabuchi M., Kishi F., Ouchi K.,
 RA Shiba T., Ishii K., Hattori M., Kihara S., Nakazawa T.;
 RA "Comparison of whole genome sequences of Chlamydia pneumoniae J138
 RT from Japan and CWL029 from USA.";
 RL Nucleic Acids Res. 28:2311-2314(2000).
 CC -1- CATALYTIC ACTIVITY: D-glyceraldehyde 3-phosphate + phosphate +
 NAD(+) = 3-phospho-D-glyceroyl phosphate + NADH.

CC -1- PATHWAY: Second phase of glycolysis; first step.
 CC -1- SUBUNIT: Homotrimer (By similarity).
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -1- SIMILARITY: Belongs to the glyceraldehyde 3-phosphate
 CC dehydrogenase family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL; AB001647; AAD18763.1; -
 CC EMBL; AE002173; AAF38006.1; -
 CC EMBL; AP002547; BAA38831.1; -
 CC PIR; B72053; B72053.
 CC PIR; E86568; E86568.
 CC HSSP; P06977; IGAD.
 CC PHCI-2DPAGE; Q92770; -
 CC TIGR; CP0123; -
 CC InterPro; IPR000173; GAP_dh_dh_dh.
 CC InterPro; IPR006424; GAPDH-I.
 CC Pfam; PF00044; spdh; 1.
 CC Pfam; PF02800; spdh; 1.
 CC PRINTS; PR00078; G3PDHGRNASE.
 CC TIGRFAMs; TIGR01534; GAPDH-I; 1.
 CC PROSITE; PS00071; GAPDH; 1.
 CC KEGG; glycolysis; Oxidoreductase; NAD; Complete proteome
 KW Glycolysis; Oxidoreductase; NAD; Complete proteome
 FT BINDING 151
 FT ACT_SITE 178 178
 FT ACTIVATES THIOL GROUP DURING CATALYSIS
 FT (BY SIMILARITY).

QY 2 NDHLSILEAMSDNTPY 18
 DB 303 NDRFFKLVAMVDNENGY 319

Query Match 38.7%; Score 46; DB 1; Length 335;
 Best Local Similarity 52.9%; Pred. No. 8.2;
 Matches 9; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

RESULT 15
 PCNA_PSEAE STANDARD; PRT; 476 AA.
 ID PCNA_PSEAE
 AC Q51508;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Salicylate biosynthesis isochorismate synthase (EC 5.4.99.6)
 DE (isochorismate mutase).
 GN PCNA OR PA4231.
 OS Pseudomonas aeruginosa.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
 OC Pseudomonadaceae; Pseudomonas.
 NCBI_TaxID=287;
 RX SEQUENCE FROM N.A.
 RP STRAIN=ATCC 15692 / PA01;
 RX MEDLINE=96086939; PubMed=7500944;
 RA Serino L., Reimann C., Baur H., Beyeler M., Viesca P., Haas D.;
 RA "Structural genes for salicylate biosynthesis from chorismate in
 RT Pseudomonas aeruginosa.";
 RL Mol. Gen. Genet. 249:217-228(1995).
 RL [2]
 RN SEQUENCE FROM N.A.
 RP STRAIN=ATCC 15692 / PA01;
 RX MEDLINE=20437337; PubMed=10984043;
 RA Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warren P.,
 RA Hickey M.J., Brinkman F.S.L., Huftagle W.O., Kowalik D.J., Lagrou M.,

RA Garber R.L., Goltzy L., Tolentino E., Westbrock-Wadman S., Yuan Y.,
 RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,
 RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
 RA Reizer J., Sailer M.H., Hancock R.E.W., Lozy S., Olson M.V.,
 RT "Complete genome sequence of *Pseudomonas aeruginosa* PA01, an
 RT opportunistic pathogen." ;
 RL Nature 406:959-964(2000).
 CC -1- FUNCTION: INVOLVED IN THE CONVERSION OF CHORISMATE TO SALICYLATE
 CC (PROBABLE).
 CC -1- CATALYTIC ACTIVITY: Chorismate = isochorismate.
 CC -1- PATHWAY: Salicylic acid biosynthesis.
 CC -1- SIMILARITY: BELONGS TO THE ISOCHORISMATE SYNTHASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; X82644; CAAS7969.1; -;
 DR EMBL; AE004840; AAG07619.1; -;
 DR PIR; S60203; S58229.
 DR InterPro: IPR005801; Anth synth chor.
 DR InterPro: IPR004561; Isochor syn.
 DR Pfam; PF00425; chorismate_bind; 1.
 DR ProDom; PD000779; Anth_synth_chor; 1.
 DR TrEMBL; TIGR00543; Isochor_syn; 1.
 KW isomerase; Complete proteome.
 SQ SEQUENCE 476 AA; 52071 MW; D0F6AFD9CDCF9CF5 CRC64;

 Query Match 38.7%; Score 46; DB 1; Length 476;
 Best Local Similarity 44.4%; Pred. No. 13;
 Matches 8; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

 QY 3 DHLSTLEAWSDDNDTPYLH 20
 DB 44 DPLOVFGAWDDDRQTPTCY 61

Search completed: November 13, 2003, 09:45:26
 Job time : 8.56398 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 13, 2003, 09:31:40 ; Search time 13.0379 Seconds
(without alignments)
154.898 Million cell updates/sec

Title: US-09-290-049a-1

Perfect score: 119

Sequence: 1 ANDHLSTLEAWSDNDPTPLMD 21

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :
1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
1	119	100.0	1475	B33135	glfB protein precu
2	110	92.4	1375	JT0345	dextranucrase (EC
3	108	90.8	1592	A38175	glucosyltransferas
4	78	65.5	1365	A41483	glucosyltransferas
5	75	63.0	1431	A45866	dextranucrase (EC
6	74	62.2	1508	T31098	probable dextranu
7	66	55.5	1449	T30857	glucosyltransferas
8	66	55.5	1449	T30852	glucosyltransferas
9	63	52.9	1577	T30858	glucosyltransferas
10	58	48.7	1518	A44811	glyceralddehyde-3-p
11	57	47.9	1331	B48445	glucosyltransferas
12	57	47.9	1599	S22737	glucosyltransferas
13	52	43.7	4848	T30289	pristinamycin 1 sy
14	50	42.0	347	T48610	hypothetical prote
15	49.5	41.6	549	T03983	rif2 nuclear resor
16	49	41.2	302	AG0597	conserved hypotet
17	49	41.2	336	QOBE40	BGMF2 protein - hu
18	49	41.2	601	E87028	pyruvate, phosphat
19	48	40.3	490	H70538	probable ppk prot
20	47.5	39.9	331	DEUTGC	glyceralddehyde-3-p
21	47.5	39.9	519	C86372	hypothetical prote
22	47	39.5	175	C86205	hypothetical prote
23	47	39.5	378	T04254	hypothetical prote
24	47	39.5	524	D82220	conserved hypotet
25	47	39.5	525	T40088	RnGEF domain cont
26	46.5	39.1	418	D90506	4-aminobutylate am
27	46	38.7	275	S55978	hypothetical prote
28	46	38.7	335	B86568	glyceralddehyde-3-p
29	46	38.7	335	B72053	glyceralddehyde 3-p

30	46	38.7	400	2	F98138	hypothetical prote
31	46	38.7	417	2	AE3149	glucosyltransferas
32	46	38.7	476	1	S58229	salicylate biosynt
33	45.5	38.2	96	2	E81786	conserved hypotet
34	45	37.8	148	2	B95908	guanydinocetate N
35	45	37.8	236	2	S62732	conserved hypotet
36	45	37.8	309	2	D95905	hypothetical prote
37	45	37.8	947	2	E86362	glycosidase homolo
38	45	37.8	1990	2	AG1749	glycosidase homolo
39	45	37.8	1091	2	AF1380	probable phosphodi
40	45	37.8	1159	2	T43461	conserved hypotet
41	44.5	37.4	210	2	H83332	transcription regu
42	44	37.0	155	2	AD2392	glyceralddehyde-3-p
43	44	37.0	247	2	PC0179	glyceralddehyde-3-p
44	44	37.0	331	1	DEECG3	glyceralddehyde-3-p
45	44	37.0	331	2	H90939	glyceralddehyde-3-p

ALIGNMENTS

RESULT 1
B33135
glfB protein precursor - Streptococcus mutans
C:Species: Streptococcus mutans
C:Date: 23-Oct-1990 #sequence_revision 23-Oct-1990 #text_change 15-Oct-1999
C:Accession: B33135; A33128
R:Shiroza, T.; Ueda, S.; Kuramitsu, H.K.
J. Bacteriol. 169, 4263-4270, 1987
A:Title: Sequence analysis of the glfB gene from Streptococcus mutans.
A:Reference number: A33135; PMID:87308013; PMID:3040685
A:Accession: B33135
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1475 <SH3>
A:Cross-references: GB:M17361; NID:G153639; PIDN:AA88588.1; PID:G153640
R:Shiroza, T.; Ueda, S.; Kuramitsu, H.K.
submitted to the Protein Sequence Database, September 1990
A:Reference number: A33128
A:Accession: A33128
A>Status: preliminary; not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 1-171,173-641,'N',643-1475 <SH2>
A:Experimental source: strain GS-5
A:Superfamily: cpl repeat homology
F:1096-1115/Domain: cpl repeat homology <CP1>
F:1224-1243/Domain: cpl repeat homology <CP2>
F:1289-1308/Domain: cpl repeat homology <CP3>
F:1354-1373/Domain: cpl repeat homology <CP4>
F:1419-1438/Domain: cpl repeat homology <CP5>

Query Match 100.0%; Score 119; DB 2; Length 1475;
Best Local Similarity 100.0%; Pred. No. 5.2e-10;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ANDHLSTLEAWSDNDPTPLMD 21
DB 481 ANDHLSTLEAWSDNDPTPLMD 501

RESULT 2
JT0345
dextranucrase (EC 2.4.1.5) precursor - Streptococcus mutans (strain GS-5)
N:Alternate names: Sucrose 6-glucosyltransferase
C:Species: Streptococcus mutans
C:Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 24-Sep-1999
C:Accession: JT0345; C3135
R:Ueda, S.; Shiroza, T.; Kuramitsu, H.K.
Gene 69, 101-109, 1988
A:Title: Sequence analysis of the glfC gene from Streptococcus mutans GS-5.
A:Reference number: JT0345; PMID:89137980; PMID:2976010
A:Accession: JT0345
A:Molecule type: DNA

A:Residues: 1-1375 <UED>
A:Experimental source: GS-5
R:Shiroza, T.; Ueda, S.; Kuramitsu, H.K.
J. Bacteriol. 169, 4263-4270, 1987
A:Title: Sequence analysis of the *gtfB* gene from *Streptococcus mutans*.
A:Reference number: A33155; MUID:87308013; PMID:3040685
A:Accession: C33135
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-349 <SHI>
A:Cross-references: GB:ML7361
C:Genetics:
A:Gene: *gtfC*
C:Function:
A:Description: catalyzes the synthesis of both water-soluble and water-insoluble glucans
C:Superfamily: cpl repeat homology
C:Keywords: duplication; glycosyltransferase; hexosyltransferase
F:1-34/Domain: signal sequence #status predicted <Sig>
F:35-1375/Product: glucosyltransferase #status predicted <Mat>
F:1126-1145/Domain: cpl repeat homology <CP1>
F:1253-1272/Domain: cpl repeat homology <CP2>
F:1318-1337/Domain: cpl repeat homology <CP3>

Query Match 92.4%; Score 110; DB 2; Length 1375;
Best Local Similarity 95.2%; Pred. No. 1,2e-08;
Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 ANDHSLTEAWSGNDPTPLHD 21
Db 507 ANDHSLTEAWSGNDPTPLHD 527

RESULT 3

A38175
glucosyltransferase precursor - *Streptococcus sobrinus*
C:Species: *Streptococcus sobrinus*
C>Date: 28-Aug-1992 #sequence_revision 28-Aug-1992 #text_change 15-Oct-1995
C:Accession: A38175
R:Abdo, H.; Mtsunura, T.; Kodama, T.; Ohta, H.; Fukui, K.; Kato, K.; Kagawa, H.
J. Bacteriol. 173, 989-996, 1991
A:Title: Peptide sequences for sucrose splitting and glucan binding within *Streptococcus*
A:Reference number: A38175; MUID:91123227; PMID:1704006
A:Accession: A38175
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1592 <ABO>
A:Cross-references: GB:D90213; NID:9217032; PIDN:BA41241.1; PID:dl014946; PID:9217033
C:Superfamily: cpl repeat homology
F:1093-1112/Domain: cpl repeat homology <CP1>
F:1222-1241/Domain: cpl repeat homology <CP2>
F:1287-1306/Domain: cpl repeat homology <CP3>
F:1330-1351/Domain: cpl repeat homology <CP4>
F:1352-1371/Domain: cpl repeat homology <CP5>
F:1402-1420/Domain: cpl repeat homology <CP6>
F:1465-1484/Domain: cpl repeat homology <CP7>
F:1513-1532/Domain: cpl repeat homology <CP8>

Query Match 90.8%; Score 108; DB 2; Length 1592;
Best Local Similarity 85.7%; Pred. No. 3e-08;
Matches 18; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

OY 1 ANDHSLTEAWSGNDPTPLHD 21
Db 477 ANNHSLTEAWSGNDPTPLHD 497

RESULT 4

A41483
glucosyltransferase (EC 2.4.1.-) *gtfs* precursor - *Streptococcus sobrinus*
C:Species: *Streptococcus sobrinus*
C>Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 15-Oct-1995
C:Accession: A41483
R:Gilmore, K.S.; Russell, R.R.B.; Ferretti, J.J.

Infect. Immun. 58, 2452-2458, 1990
A:Title: Analysis of the *Streptococcus downei* *gtfs* gene, which specifies a glucosyltransferase
A:Reference number: A41483; MUID:90316665; PMID:2142479
A:Accession: A41483
A:Molecule type: DNA
A:Residues: 1-1365 <GIL>
A:Cross-references: GB:M30943; NID:9153652; PIDN:AAA26898.1; PID:9153653
C:Genetics:
A:Gene: *gtfs*
C:Superfamily: cpl repeat homology
C:Keywords: glycosyltransferase; hexosyltransferase
Query Match 65.5%; Score 78; DB 2; Length 1365;
Best Local Similarity 71.4%; Pred. No. 0.0012;
Matches 15; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

RESULT 5

A45866
dextranucrase (EC 2.4.1.5) precursor - *Streptococcus mutans*
C:Species: *Streptococcus mutans*
C>Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 07-May-1999
C:Accession: A45866
R:Honda, O.; Kato, C.; Kuramitsu, H.K.
J. Gen. Microbiol. 136, 2099-2105, 1990
A:Title: Nucleotide sequence of the *Streptococcus mutans* *gtfd* gene encoding the glucosyl
A:Reference number: A45866; MUID:91100958; PMID:2148600
A:Accession: A45866
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1431 <HON>
A:Cross-references: GB:M29296
C:Superfamily: cpl repeat homology
C:Keywords: glycosyltransferase; hexosyltransferase
F:181-201/Domain: cpl repeat homology <CP1>
F:1127-1146/Domain: cpl repeat homology <CP2>
F:1192-1211/Domain: cpl repeat homology <CP3>
F:1257-1276/Domain: cpl repeat homology <CP4>
F:1277-1297/Domain: cpl repeat homology <CP5>
F:1321-1340/Domain: cpl repeat homology <CP6>
F:1341-1361/Domain: cpl repeat homology <CP7>
F:1385-1404/Domain: cpl repeat homology <CP8>

Query Match 63.0%; Score 75; DB 2; Length 1431;
Best Local Similarity 71.4%; Pred. No. 0.0036;
Matches 15; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

OY 1 ANDHSLTEAWSGNDPTPLHD 21
Db 495 ANNHSLTEAWSGNDPTPLHD 515

RESULT 6

T31098
probable dextranucrase (EC 2.4.1.5), extracellular - *Leuconostoc mesenteroides*
C:Species: *Leuconostoc mesenteroides*
C>Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 11-May-2000
C:Accession: T31098
R:Monchois, V.; Renaud-Simeon, M.; Monsan, P.; Willemot, R.M.
FEMS Microbiol. Lett. 159, 307-315, 1998
A:Title: Cloning and sequencing of a gene coding for an extracellular dextranucrase (DS
A:Reference number: Z20981; MUID:98164374; PMID:9503626
A:Accession: T31098
A:Status: preliminary; translated from GB/EMBL/DBD
A:Molecule type: DNA
A:Residues: 1-1508 <MON>
A:Cross-references: EMBL:AF030129; NID:92766611; PID:92766612; PIDN:AA95453.1
A:Experimental source: strain NRRL B-1299
C:Genetics:

A:Gene: dsrB
A:Function:
A:Description: produces dextran composed only of alpha(1-6) glucosidic bonds
C:Keywords: glycosyltransferase, hexosyltransferase

Query Match 62.2%; Score 74; DB 2; Length 1508;
Best Local Similarity 66.7%; Pred. No. 0.0055;
Matches 14; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 1 ANDHSLLEAMSDNDTPYLHD 21
DB 563 ANOHLHLEDMSHNDPEYKD 583

RESULT 7
T30857
glucosyltransferase - Streptococcus salivarius
C:Species: Streptococcus salivarius
C:Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 22-Oct-1999
C:Accession: T30857
R:Simpson, C.L.; Giffard, P.M.; Jacques, N.A.
Infect. Immun. 63, 609-621, 1995
A:Title: Streptococcus salivarius ATCC 25975 possesses at least two genes coding for pri
A:Reference number: Z20909; MUID:95122197; PMID:7822030
A:Accession: T30857
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1449 <SIM>
A:Cross-references: EMBL:L35495; NID:g662378; PID:g662379; PIDN:AA041412.1
C:Keywords:
A:Gene: gtfI

Query Match 55.5%; Score 66; DB 2; Length 1449;
Best Local Similarity 72.2%; Pred. No. 0.092;
Matches 13; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 4 HSLLEAMSDNDTPYLHD 21
DB 539 HSLLEAMSHNDAYNED 556

RESULT 8
T30552
glucosyltransferase N - Streptococcus salivarius (fragment)
C:Species: Streptococcus salivarius
C:Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 22-Oct-1999
C:Accession: T30552
R:Jaife, R.I.
submitted to the EMBL Data Library, February 1998
A:Description: Streptococcus salivarius V1477 gtfN.
A:Reference number: Z20854
A:Accession: T30552
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1449 <JAF>
A:Cross-references: EMBL:AF049609; NID:g2935545; PID:g2935546; PIDN:AA05155.1
C:Keywords:
A:Gene: gtfN

Query Match 55.5%; Score 66; DB 2; Length 1449;
Best Local Similarity 72.2%; Pred. No. 0.092;
Matches 13; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 4 HSLLEAMSDNDTPYLHD 21
DB 539 HSLLEAMSHNDAYNED 556

RESULT 9
T30858
glucosyltransferase - Streptococcus salivarius
C:Species: Streptococcus salivarius
C:Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 22-Oct-1999

C:Accession: T30858
R:Simpson, C.L.; Giffard, P.M.; Jacques, N.A.
Infect. Immun. 63, 609-621, 1995
A:Title: Streptococcus salivarius ATCC 25975 possesses at least two genes coding for pri
A:Reference number: Z20909; MUID:95122197; PMID:7822030
A:Accession: T30858
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1577 <SIM>
A:Cross-references: EMBL:L35928; NID:g662380; PID:g662381; PIDN:AA041413.1
C:Keywords:
A:Gene: gtfM

Query Match 52.9%; Score 63; DB 2; Length 1577;
Best Local Similarity 72.2%; Pred. No. 0.3;
Matches 13; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 4 HSLLEAMSDNDTPYLHD 21
DB 591 HSLLEAMSYNDHOYKND 608

RESULT 10
A44811
glucosyltransferase (EC 2.4.1.-) I - Streptococcus salivarius
C:Species: Streptococcus salivarius
C:Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 15-Oct-1999
C:Accession: A44811; S22726; S28809
R:Giffard, P.M.; Simpson, C.L.; Milward, C.P.; Jacques, N.A.
J. Gen. Microbiol. 137, 2577-2593, 1991
A:Title: Molecular characterization of a cluster of at least two glucosyltransferase gene
A:Reference number: A44811; MUID:92148377; PMID:1838391
A:Accession: A44811
A:Molecule type: DNA
A:Residues: 1-1518 <GIF>
A:Cross-references: EMBL:Z11873; NID:g47526; PIDN:CAA77900.1; PID:g47527
A:Note: Sequence extracted from NCBI backbone (NCBIN:81050, NCBI:P:81052)
C:Keywords:
A:Gene: gtfJ
A:Superfamily: cpl repeat homology
C:Keywords: glycosyltransferase; hexosyltransferase
P:1307-1326/Domain: cpl repeat homology <CP4>

Query Match 48.7%; Score 58; DB 2; Length 1518;
Best Local Similarity 66.7%; Pred. No. 1.7;
Matches 10; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 4 HSLLEAMSDNDTPY 18
DB 504 HSLLEAMSLNDNH 518

RESULT 11
B48445
glyceralddehyde-3-phosphate dehydrogenase (phosphorylating) (EC 1.2.1.12) - Leishmania mex
C:Species: Leishmania mexicana
C:Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 03-Jun-2002
C:Accession: B48445; S25142
R:Hannaert, V.; Blaauw, M.; Kohl, L.; Allert, S.; Opperdoes, F.R.; Michels, P.A.M.
Mol. Biochem. Parasitol. 55, 115-126, 1992
A:Title: Molecular analysis of the cytosolic and glycosomal glyceralddehyde-3-phosphate de
A:Reference number: A48445; MUID:93063042; PMID:1435864
A:Accession: B48445
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-331 <HAN>
A:Cross-references: EMBL:X65220; NID:g9552; PIDN:CAA46323.1; PID:g9553
C:Superfamily: glyceralddehyde-3-phosphate dehydrogenase
C:Keywords: oxidoreductase

Query Match 47.9%; Score 57; DB 2; Length 331;
Best Local Similarity 52.6%; Pred. No. 0.4;
Matches 10; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 2 NDHSLLEAWSNDNTPYLH 20
DB 301 NDHFVKLWSWYDNETGYSH 319

RESULT 12

glucosyltransferase (EC 2.4.1.-) S - Streptococcus salivarius
C/Species: Streptococcus salivarius
C/Date: 16-Sep-1992 #sequence_revision 16-Sep-1992 #text_change 21-Jan-2000
C/Accession: S22737; S28810; B44811; S22727
R/Jacques, N.
submitted to the EMBL Data Library, March 1992
A/Reference number: S22726
A/Accession: S22737
A/Molecule type: DNA
A/Residues: 1-1599 <UAC>
A/Cross-references: EMBL:Z11872; NID:G47530; PIDN:CA77898.1; PID:G47531
A/Experimental source: ATCC 25975
R/Giffard, P.M.; Simpson, C.L.; Milward, C.P.; Jacques, N.A.
J. Gen. Microbiol. 137, 2577-2593, 1991
A/Title: Molecular characterization of a cluster of at least two glucosyltransferase genes
A/Reference number: A44811; MUID:92148377; PMID:1838391
A/Accession: S28810
A/Molecule type: DNA
A/Residues: 1-51 <GIF>
A/Cross-references: EMBL:Z11873
C/Genetics:
A/Status: preliminary
A/Note: rft2
C/Keywords: glycosyltransferase; hexosyltransferase
F:1456-1475/Domain: cpl repeat homology <CPR>

Query Match 47.9%; Score 57; DB 2; Length 1599;
Best Local Similarity 55.6%; Pred. No. 2.6;
Matches 10; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

QY 4 HSLLEAWSNDNTPYLHD 21
DB 494 NLSLEAWSNDNTPYVNE 511

RESULT 13
T30289
pristinamycin I synthase 3 - Streptomyces pristinaespiralis
C/Species: Streptomyces pristinaespiralis
C/Date: 02-Sep-2000 #sequence_revision 02-Sep-2000 #text_change 01-Dec-2000
C/Accession: T30289
R/Creevy-Lagard, V.A.; Saurin, W.; Thibaut, D.; Gil, P.; Naudin, L.; Crouzet, J.; Bian
submitted to the EMBL Data Library, February 1997
A/Description: Streptogramin B biosynthesis in Streptomyces pristinaespiralis and Strept
A/Reference number: Z20808
A/Accession: T30289
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-4848 <DEC>
A/Cross-references: EMBL:Y11548; NID:e1025755; PID:e307539; PIDN:CA72312.1
C/Genetics:
A/Status: rft2
C/Keywords: carrier protein; phosphopantetheine; phosphoprotein
F:511-951/Domain: acetate-CoA ligase homology <ACL1>
F:966-1034/Domain: acyl carrier protein homology <ACPP>
F:1563-2024/Domain: acetate-CoA ligase homology #status atypical <ACL2>
F:2399-2467/Domain: acyl carrier protein homology <ACP2>
F:2995-3427/Domain: acetate-CoA ligase homology <ACL3>
F:3441-3509/Domain: acyl carrier protein homology <ACP3>
F:4043-4492/Domain: acetate-CoA ligase homology <ACL4>
F:4507-4575/Domain: acyl carrier protein homology <ACP4>
F:998,2431,3473/Binding site: phosphopantetheine (Ser) (covalent) #status predicted

Query Match 43.7%; Score 52; DB 2; Length 4848;

Best Local Similarity 47.1%; Pred. No. 58;
Matches 8; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 1 ANDHSLLEAWSNDNTP 17
DB 3972 ADHARVLGWNDDTDP 3988

RESULT 14

T48610
hypothetical protein F18022.160 - Arabidopsis thaliana
C/Species: Arabidopsis thaliana (mouse-ear cress)
C/Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 20-Apr-2000
C/Accession: T48610
R/Bevan, M.; Murphy, G.; Ridley, P.; Hudson, S.; Bancroft, I.; Mewes, H.W.; Rudd, S.; Le
submitted to the Protein Sequence Database, April 2000
A/Reference number: Z24493
A/Accession: T48610
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-347 <BEV>
A/Cross-references: EMBL:DB
A/Experimental source: cultivar Columbia, BAC clone F18022
C/Genetics:
A/Map position: 5
A/Intons: 232/3; 296/3
A/Note: F18022.160

Query Match 42.0%; Score 50; DB 2; Length 347;
Best Local Similarity 52.9%; Pred. No. 5.2;
Matches 9; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 3 DHSLEAWSNDNTPYL 19
DB 200 DYEQIMEAWSDKGLYV 216

RESULT 15

T03983
rf2 nuclear restorer protein - maize
N/Alternate names: aldehyde dehydrogenase homolog
C/Species: Zea mays (maize)
C/Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 11-May-2000
C/Accession: T03983
R/Cui, X.; Wise, R.P.; Schnable, P.S.
Science 272, 1334-1336, 1996
A/Title: The rf2 nuclear restorer gene of male-sterile T-cytoplasm maize.
A/Reference number: Z15177; MUID:96243131; PMID:8650543
A/Accession: T03983
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: mRNA
A/Residues: 1-549 <CUI>
A/Cross-references: EMBL:U43082; NID:g1421729; PIDN:AAC49371.1; PID:g1421730
A/Experimental source: strain B73
C/Genetics:
A/Map position: 9
A/Note: rft2
C/Keywords: aldehyde dehydrogenase (NAD+); aldehyde dehydrogenase homology
F:107-370/Domain: aldehyde dehydrogenase homology <ALD>

Query Match 41.6%; Score 49.5; DB 2; Length 549;
Best Local Similarity 58.8%; Pred. No. 11;
Matches 10; Conservative 1; Mismatches 5; Indels 1; Gaps 1;

QY 2 NDHSLLEAWSNDNTPY 18
DB 147 NDHALALFTW-DNGKPY 162

Search completed: November 13, 2003, 09:50:21
Job time: 15.0379 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 13, 2003, 09:31:40 ; Search time 31.6493 Seconds
(without alignments)
171.224 Million cell updates/sec

Title: US-09-290-049a-1
Perfect score: 119
Sequence: 1 ANDHSLTEAWSNDTPYLHD 21

Scoring table:
BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteriaph:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	108	90.8	1590	2	Q59983 streptococc
2	108	90.8	1590	2	Q55263 streptococc
3	78	65.5	1338	2	Q9WJ04 streptococc
4	76	63.9	1016	2	Q9LJ07 leuconostoc
5	74	62.2	1477	2	Q9L466 leuconostoc
6	74	62.2	1508	2	Q9E2H5 leuconostoc
7	74	62.2	1508	2	Q52224 leuconostoc
8	73	61.3	1575	2	Q9LJ03 streptococc
9	73	61.3	1577	2	Q54178 streptococc
10	71	59.7	1527	2	Q9ZAR4 leuconostoc
11	71	59.7	1527	2	Q8KRE1 leuconostoc
12	69	58.0	2835	2	Q8G9Q2 leuconostoc
13	66	55.5	1449	2	Q68542 streptococc
14	66	55.5	1449	2	Q55264 streptococc
15	63	52.9	1554	2	Q8KZL5 streptococc
16	63	52.9	1577	2	Q55265 streptococc

17	62	52.1	1512	2	Q9WJ05	Q9WJ05 streptococc
18	60	50.4	2057	2	Q9RE05	Q9RE05 leuconostoc
19	58	48.7	1518	2	Q00600	Q00600 streptococc
20	57	47.9	1599	2	Q00599	Q00599 streptococc
21	55.5	46.6	282	10	Q8RW33	Q8RW33 allium cepa
22	52	43.7	4848	2	Q07944	Q07944 streptomyces
23	51	42.9	424	10	Q94J07	Q94J07 arabidopsis
24	51	42.9	435	10	Q9LJ08	Q9LJ08 arabidopsis
25	50	42.0	305	10	Q8LDL1	Q8LDL1 arabidopsis
26	50	42.0	339	10	Q94979	Q94979 arabidopsis
27	50	42.0	347	10	Q9LY92	Q9LY92 arabidopsis
28	49.5	41.6	55	10	Q94FX9	Q94FX9 zea mays (m
29	48.5	41.6	55	10	Q93VW4	Q93VW4 zea mays (m
30	48.5	41.6	55	10	Q93WH7	Q93WH7 zea mays (m
31	48.5	41.6	56	10	Q94FX8	Q94FX8 zea mays (m
32	48.5	41.6	129	10	Q94FX6	Q94FX6 zea mays (m
33	48.5	41.6	547	10	Q8LST4	Q8LST4 sorghum bic
34	49.5	41.6	549	10	Q94G64	Q94G64 zea mays (m
35	49.5	41.6	549	10	Q43274	Q43274 zea mays (m
36	49	41.2	601	16	Q05566	Q05566 mycobacteri
37	49	41.2	797	16	Q98B32	Q98B32 rhizobium l
38	48.5	40.8	283	5	Q9V1P7	Q9V1P7 drosophila
39	48.5	40.8	283	5	Q8WSF4	Q8WSF4 drosophila
40	48	40.3	490	15	Q06579	Q06579 mycobacteri
41	47.5	39.9	214	11	Q8C1P7	Q8C1P7 mus musculu
42	47.5	39.9	519	10	Q9ZUB6	Q9ZUB6 arabidopsis
43	47.5	39.9	534	10	Q8S528	Q8S528 arabidopsis
44	47.5	39.9	534	10	Q94C67	Q94C67 arabidopsis
45	47.5	39.9	549	10	Q9LIR2	Q9LIR2 oryza sativ

ALIGNMENTS

RESULT 1

Q59983 PRELIMINARY; PRT; 1590 AA.

AC Q59983; 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-OCT-2002 (TREMBLrel. 22, Last annotation update)
DE Glucosyltransferase-I precursor (EC 2.4.1.5).
GN GTFI.
OS Streptococcus sobrinus.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1310;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=QWZ176;
RX MEDLINE=94146405; PubMed=8312602;
RA Sato S., Inoue M., Handa N., Aizawa Y., Isobe Y., Katayama T.;
RT "DNA sequence of the glucosyltransferase gene of serotype d
RT Streptococcus sobrinus.";
RL DNA Seq. 4:19-27(1993).
DR EMBL; D13858; BAA02976.1; -;
DR InterPro; IPR002478; CW_binding.
DR InterPro; IPR003318; Glyco_hydro_70.
DR Pfam; PF01473; CW_binding_1; 16.
DR Pfam; PF02324; Glyco_hydro_70; 1.
KW GLYCOSYLTRANSFERASE; Signal; Transferase.
FT SIGNAL 1
FT CHAIN 39 1590
FT SEQUENCE 1590 AA; 175955 MW; C3C83A57CF3C2B08 CRC64;

Query Match 90.8%; Score 108; DB 2; Length 1590;
Best Local Similarity 85.7%; Pred. No. 5.5e-08;
Matches 18; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 ANDHSLTEAWSNDTPYLHD 21
Db 477 ANNHSLTEAWSNDTPYLHD 497

RESULT 2

Q55263 PRELIMINARY; PRT; 1590 AA.
 ID Q55263
 AC Q55263;
 DT 01-NOV-1996 (TREMBLrel. 01, Created)
 DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
 DT 01-OCT-2002 (TREMBLrel. 22, Last annotation update)
 DE GTF-1.
 GN GLUCOSYLTRANSFERASE.
 OS Streptococcus sobrinus.
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 CC Streptococcus.
 OX NCBI_TaxID=1310;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 33478;
 RA Sato S.;
 RT "DNA and amino-acid sequences of water-insoluble-glucan synthetase produced from Streptococcus sobrinus ATCC 33478.";
 RL Am. Kagoshima Univ. Dental School 16:23-29(1996).
 DR EMBL; D63570; BAA09792.1; -
 DR InterPro; IPR002479; CW_binding.
 DR InterPro; IPR003318; Glyco_hydro_70.
 DR Pfam; PF01473; CW_binding_1; 15.
 DR Pfam; PF02324; Glyco_hydro_70; 1.
 DR SEQUENCE 1590 AA; 176058 MW; 9DF7A3F2C6E4FD43 CRC64;

Query Match 90.8%; Score 108; DB 2; Length 1590;
 Best Local Similarity 85.7%; Pred. No. 5.5e+08;
 Matches 18; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 ANDHSLILEAWSDNDTPYLHD 21
 DB 477 ANNHVSIVEAWSDNDTPYLHD 497

RESULT 3

Q9WKXJ4 PRELIMINARY; PRT; 1338 AA.
 ID Q9WKXJ4
 AC Q9WKXJ4;
 DT 01-NOV-1999 (TREMBLrel. 12, Created)
 DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
 DT 01-OCT-2002 (TREMBLrel. 22, Last annotation update)
 DE GTF-8.
 GN Streptococcus criceti.
 OS Streptococcus criceti.
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 CC Streptococcus.
 OX NCBI_TaxID=1333;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=HS-6;
 RA Inoue M., Fukui K., Miyagi A.;
 RT "S. cricetus glucosyltransferase (gfts and gtf8) genes.";
 RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AB026123; BAA77236.1; -
 DR InterPro; IPR002479; CW_binding.
 DR InterPro; IPR003318; Glyco_hydro_70.
 DR Pfam; PF01473; CW_binding_1; 10.
 DR Pfam; PF02324; Glyco_hydro_70; 1.
 DR KJ plasmid.
 SQ SEQUENCE 1338 AA; 148558 MW; 0A90C8E10E15D99B CRC64;

Query Match 65.5%; Score 78; DB 2; Length 1338;
 Best Local Similarity 71.4%; Pred. No. 0.0025;
 Matches 15; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 ANDHSLILEAWSDNDTPYLHD 21
 DB 437 AIDHLSILEAWSGNDNDYKD 457

RESULT 4

Q9LCJ7 PRELIMINARY; PRT; 1016 AA.
 ID Q9LCJ7
 AC Q9LCJ7;
 DT 01-OCT-2000 (TREMBLrel. 15, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
 DT 01-OCT-2002 (TREMBLrel. 22, Last annotation update)
 DE Dextranucrase.
 GN DSRT.
 OS Leuconostoc mesenteroides.
 OC Bacteria; Firmicutes; Lactobacillales; Leuconostoc.
 OX NCBI_TaxID=1245;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=NRRL B-512F;
 RX MEDLINE=20169623; PubMed=10705445;
 RA Punane K., Mizuno K., Takahara H., Kobayashi M.;
 RT "Gene encoding a dextranucrase-like protein in Leuconostoc mesenteroides NRRL B-512F.";
 RL Biosci. Biotechnol. Biochem. 64:29-38(2000).
 DR EMBL; AB020020; BAA90527.1; -
 DR HSSP; P06278; IVS.
 DR InterPro; IPR003318; Glyco_hydro_70.
 DR Pfam; PF02324; Glyco_hydro_70; 1.
 DR SEQUENCE 1016 AA; 110344 MW; 8896EPD13CCCB47 CRC64;

Query Match 63.9%; Score 76; DB 2; Length 1016;
 Best Local Similarity 61.9%; Pred. No. 0.0037;
 Matches 13; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 ANDHSLILEAWSDNDTPYLHD 21
 DB 548 SNOCHSILBDSNDNDPEYKD 568

RESULT 5

Q9L466 PRELIMINARY; PRT; 1477 AA.
 ID Q9L466
 AC Q9L466;
 DT 01-OCT-2000 (TREMBLrel. 15, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
 DT 01-OCT-2002 (TREMBLrel. 22, Last annotation update)
 DE Dextranucrase (EC 2.4.1.5).
 GN DSRC.
 OS Leuconostoc mesenteroides.
 OC Bacteria; Firmicutes; Lactobacillales; Leuconostoc.
 OX NCBI_TaxID=1245;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=NRRL B-1355;
 RA Arguello-Morales M.A., Remaud-Simeon M., Pitzut S., Sarcabal P., Willemot R.M., Monsan P.;
 RT "Sequence analysis of the gene encoding alternanucrase, a sucrose glucosyltransferase from Leuconostoc mesenteroides NRRL B-1355.";
 RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AJ250172; CAB76565.1; -
 DR InterPro; IPR002479; CW_binding.
 DR InterPro; IPR003318; Glyco_hydro_70.
 DR Pfam; PF01473; CW_binding_1; 14.
 DR Pfam; PF02324; Glyco_hydro_70; 1.
 DR KX Glycosyltransferase; Transferase.
 SQ SEQUENCE 1477 AA; 164887 MW; E6F5710DEDFCB831 CRC64;

Query Match 62.2%; Score 74; DB 2; Length 1477;
 Best Local Similarity 66.7%; Pred. No. 0.012;
 Matches 14; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 1 ANDHSLILEAWSDNDTPYLHD 21
 DB 532 ANQHSLILEAWSHNDPEYKD 552

```

RESULT 6
Q9EZHS PRELIMINARY; PRT: 1508 AA.
ID Q9EZHS;
AC Q9EZHS;
DT 01-MAR-2001 (TRENBLrel. 16, Created)
DT 01-MAR-2001 (TRENBLrel. 16, Last sequence update)
DT 01-OCT-2002 (TRENBLrel. 22, Last annotation update)
DE Dextranucrase Dsb742.
GN DSRB742.
OS Leuconostoc mesenteroides.
OC Bacteria; Firmicutes; Lactobacillales; Leuconostoc.
OX NCBI_TaxID=1245;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=B-742CB;
RA Kim H.-S., Kim D., Ryu H.-J., Robyt J.F.;
RT "Leuconostoc mesenteroides B-742CB, a dextranucrase gene.";
RL Submitted (Aug-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF294469; AAG38021.1;
DR InterPro; IPR002479; CW_binding.
DR InterPro; IPR003318; Glyco_hydro_70.
DR Pfam; PF01473; CW_binding_1; 14.
DR Pfam; PF02324; Glyco_hydro_70; 1.
SQ SEQUENCE 1508 AA; 168542 MW; E2FCFA0F87AE4F3A CRC64;

Query Match 62.2%; Score 74; DB 2; Length 1508;
Best Local Similarity 66.7%; Pred. No. 0.012;
Matches 14; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 1 ANDHSTLEAWSNDPTPYLHD 21
Db 563 ANQHSLIEDMSHNDPEYKD 583

RESULT 7
Q52224 PRELIMINARY; PRT: 1508 AA.
ID Q52224;
AC Q52224;
DT 01-JUN-1998 (TRENBLrel. 06, Created)
DT 01-JUN-1998 (TRENBLrel. 06, Last sequence update)
DT 01-OCT-2002 (TRENBLrel. 22, Last annotation update)
DE Glucosyltransferase (EC 2.4.1.5).
GN DSRB.
OS Leuconostoc mesenteroides.
OC Bacteria; Firmicutes; Lactobacillales; Leuconostoc.
OX NCBI_TaxID=1245;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NRRL B-1299;
RA Monchois V., Remaud-Simeon M., Monsan P., Willemot R.M.;
RT "Cloning and sequencing of a gene coding for an extracellular
RT dextranucrase (DSRB) from Leuconostoc mesenteroides NRRL B-1299
RT synthesizing only a a(1-6) glucan.";
RL FEMS Microbiol. Lett. 0:0-0(1998).
DR EMBL; AF030129; AAB95453.1;
DR InterPro; IPR002479; CW_binding.
DR InterPro; IPR003318; Glyco_hydro_70.
DR Pfam; PF01473; CW_binding_1; 14.
DR Pfam; PF02324; Glyco_hydro_70; 1.
KM Glycosyltransferase;
SQ SEQUENCE 1508 AA; 168511 MW; E70CECB57A70D1F0 CRC64;

Query Match 62.2%; Score 74; DB 2; Length 1508;
Best Local Similarity 66.7%; Pred. No. 0.012;
Matches 14; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 1 ANDHSTLEAWSNDPTPYLHD 21
Db 563 ANQHSLIEDMSHNDPEYKD 583

RESULT 8
Q9LCH3 PRELIMINARY; PRT: 1575 AA.
ID Q9LCH3;
AC Q9LCH3;
DT 01-OCT-2000 (TRENBLrel. 15, Created)
DT 01-OCT-2000 (TRENBLrel. 15, Last sequence update)
DT 01-OCT-2002 (TRENBLrel. 22, Last annotation update)
DE Glucosyltransferase.
GN GTRF.
OS Streptococcus oralis.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1303;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC10557;
RA Fujitawa T., Hoshino T., Ooshima T., Sobue S., Hamada S.;
RT "Purification, characterization, and molecular analysis of the gene
RT encoding glucosyltransferase from Streptococcus oralis.";
RL Infect. Immun. 68:2475-2483(2000).
DR EMBL; AB025228; BAA95201.1;
DR InterPro; IPR002479; CW_binding.
DR InterPro; IPR003318; Glyco_hydro_70.
DR Pfam; PF01473; CW_binding_1; 17.
DR Pfam; PF02324; Glyco_hydro_70; 1.
KM Transferase.
SQ SEQUENCE 1575 AA; 176792 MW; 772A26E4D7C2E543 CRC64;

Query Match 61.3%; Score 73; DB 2; Length 1575;
Best Local Similarity 77.8%; Pred. No. 0.018;
Matches 14; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 4 HSLIEAWSNDPTPYLHD 21
Db 549 HSLIEAWSNDPTPYKMD 566

RESULT 9
Q54178 PRELIMINARY; PRT: 1577 AA.
ID Q54178;
AC Q54178; Q54247;
DT 01-NOV-1996 (TRENBLrel. 01, Created)
DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)
DT 01-OCT-2002 (TRENBLrel. 22, Last annotation update)
DE Glucosyltransferase.
GN GTRG.
OS Streptococcus gordonii Challis.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=29390;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CHALLIS;
RA Wickerman M.M., Sulavik M.C., Clewell D.B.;
RT "Molecular analysis of Streptococcus gordonii glucosyltransferase
RT phase variants.";
RL Dev. Biol. Stand. 85:309-314(1995).
RN [2]
RP SEQUENCE OF 1-96 FROM N.A.
RC STRAIN=CHALLIS;
RX MEDLINE=92276337; PubMed=1534326;
RA Sulavik M.C., Tardif G., Clewell D.B.;
RT "Identification of a gene, rgg, which regulates expression of
RT glucosyltransferase and influences the spp phenotype of Streptococcus
RT gordonii Challis.";
RL J. Bacteriol. 174:3577-3586(1992).
DR EMBL; U12643; AAC43483.1;
DR EMBL; M89763; AAA26969.1;
DR InterPro; IPR002479; CW_binding.
DR InterPro; IPR003318; Glyco_hydro_70.
DR Pfam; PF01473; CW_binding_1; 18.
DR Pfam; PF02324; Glyco_hydro_70; 1.
KM Transferase.

```

SQ SEQUENCE 1577 AA; 177805 MM; 5AB0328DC5E08D18 CRC64;

Query Match 61.3%; Score 73; DB 2; Length 1577;

Best Local Similarity 77.8%; Pred. No. 0.018; Mismatches 4; Indels 0; Gaps 0;

Matches 14; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 4 HSLILEAMSDNDPTPLHD 21

DB 551 HSLILEAMSDNDPTPLHD 568

RESULT 10

ID Q9ZAR4 PRELIMINARY; PRT; 1527 AA.

AC Q9ZAR4;

DT 01-MAY-1999 (TREMBlrel. 10, Created)

DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)

DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)

DE Dextranucrase.

GN DEX.

OS Leuconostoc mesenteroides.

OC Bacteria; Firmicutes; Lactobacillales; Leuconostoc.

OX NCBI_TaxID=1245;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=NRRL B-512-F;

RA Bhargava R., Singh D.K.S.;

RT "Cloning and Molecular Characterization of Dextranucrase Gene from

Leuconostoc mesenteroides NRRL B-512F."

RL Submitted (DEC-1996) to the EMBL/GenBank/DBJ databases.

DR InterPro: IPR002479; Glyco_hydro_70.

DR Pfam: PF01473; Glyco_hydro_70; 16.

DR Pfam: PF03324; Glyco_hydro_70; 1.

SQ SEQUENCE 1527 AA; 169705 MM; 1DFAFA237C743398 CRC64;

QY Query Match 59.7%; Score 71; DB 2; Length 1527;

Best Local Similarity 66.7%; Pred. No. 0.037;

Matches 14; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 1 ANDHSLILEAMSDNDPTPLHD 21

DB 581 ANOHLSTLEDMWSDNDPTPLHD 601

RESULT 11

ID Q8KRE1 PRELIMINARY; PRT; 1527 AA.

AC Q8KRE1;

DT 01-OCT-2002 (TREMBlrel. 22, Created)

DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)

DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)

DE Dextranucrase DsrD (EC 2.4.1.5).

GN DSRD.

OS Leuconostoc mesenteroides.

OC Bacteria; Firmicutes; Lactobacillales; Leuconostoc.

OX NCBI_TaxID=1245;

RN [1]

RP SEQUENCE FROM N.A.

RC "Isolation and characterization of the dextranucrase DsrD of

Leuconostoc mesenteroides Icc4."

RT Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL: AF017384; AAG61158.1; -

DR InterPro: IPR002479; Glyco_hydro_70.

DR Pfam: PF01473; Glyco_hydro_70; 6.

DR Pfam: PF03324; Glyco_hydro_70; 1.

KM Transferrase; Glycosyltransferase.

SQ SEQUENCE 1527 AA; 169835 MM; F9D0DE220BDB9668 CRC64;

QY Query Match 59.7%; Score 71; DB 2; Length 1527;

Best Local Similarity 66.7%; Pred. No. 0.037;

Matches 14; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 1 ANDHSLILEAMSDNDPTPLHD 21

DB 581 ANOHLSTLEDMWSDNDPTPLHD 601

RESULT 12

ID Q8G9Q2 PRELIMINARY; PRT; 2835 AA.

AC Q8G9Q2;

DT 01-MAR-2003 (TREMBlrel. 23, Created)

DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)

DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)

DE Dextranucrase (EC 2.4.1.5) (Fragment).

GN DSRD.

OS Leuconostoc mesenteroides.

OC Bacteria; Firmicutes; Lactobacillales; Leuconostoc.

OX NCBI_TaxID=1245;

RN [1]

RP SEQUENCE FROM N.A.

RC MEDLINE=2231661; PubMed=12270834;

RA Bozonnet S., Dolis-Lafargue M., Fabre E., Pizzut S., Renaud-Simeon M.,

Mondan P., Willemot R.M.;

RT "Molecular characterization of DSR-E, an alpha-1,2 linkage

synthesizing dextranucrase with two catalytic domains."

RU EMBL: AJ430204; CAD22883.1; -

DR Transferrase; Glycosyltransferase.

FT NON TER 1

SQ SEQUENCE 2835 AA; 313264 MM; D03262CDD735399D CRC64;

QY Query Match 58.0%; Score 69; DB 2; Length 2835;

Best Local Similarity 57.9%; Pred. No. 0.16;

Matches 11; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

QY 1 ANDHSLILEAMSDNDPTPLHD 19

DB 557 ANHSLILEDMWSDNDPTPLHD 575

RESULT 13

ID Q68542 PRELIMINARY; PRT; 1449 AA.

AC Q68542;

DT 01-AUG-1998 (TREMBlrel. 07, Created)

DT 01-AUG-1998 (TREMBlrel. 07, Last sequence update)

DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)

DE Glucosyltransferase N (Fragment).

GN GTRN.

OS Streptococcus salivarius.

OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;

OC Streptococcus.

OX NCBI_TaxID=1304;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=V1477;

RA Jaffe R.I.;

RT "Streptococcus salivarius V1477 gtfN";

RT Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.

DR EMBL: AF049609; AAC05156.1; -

DR InterPro: IPR002479; Glyco_hydro_70.

DR InterPro: IPR003318; Glyco_hydro_70.

DR Pfam: PF01473; Glyco_hydro_70; 8.

DR Pfam: PF03324; Glyco_hydro_70; 1.

KM Transferrase.

FT NON TER 1449

SQ SEQUENCE 1449 AA; 159895 MM; 0700F6D748471BFB CRC64;

QY Query Match 55.5%; Score 66; DB 2; Length 1449;

Best Local Similarity 72.2%; Pred. No. 0.21;

Matches 13; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 4 HSLILEAMSDNDTPYLHD 21
 DB 539 HSLILEAMSHNDAYNED 556

RESULT 14

Q55264 PRELIMINARY; PRT; 1449 AA.
 AC Q55264;
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-NOV-1996 (TrEMBLrel. 01, last sequence update)
 DT 01-OCT-2002 (TrEMBLrel. 22, last annotation update)
 DE Glucosyltransferase precursor.
 GN GTFU.
 OS Streptococcus salivarius.
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 OX NCBI_TaxID=1304;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=95122197; PubMed=7822030;
 RA Simpson C.L., Giffard P.M., Jacques N.A.;
 RT "Streptococcus salivarius ATCC 25975 possesses at least two genes
 coding for primer-independent glucosyltransferases.";
 RL Infect. Immun. 63:609-621 (1995).
 DR EMBL; J35495; AAC4142.1;
 DR InterPro; IPR002479; CW_binding.
 DR InterPro; IPR003318; Glyco_hydro_70.
 DR Pfam; PF01473; CW_binding_1; 8.
 DR Pfam; PF02324; Glyco_hydro_70; 1.
 KW Signal; Transferase.
 FT SIGNAL 1 35 POTENTIAL.
 FT CHAIN 36 1449 GLUCOSYLTRANSFERASE.
 SQ SEQUENCE 1449 AA; 159984 MM; DD62F07306E86A46 CRC64;

Query Match 55.5%; Score 66; DB 2; Length 1449;
 Best Local Similarity 72.2%; Pred. No. 0.21;
 Matches 13; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 4 HSLILEAMSDNDTPYLHD 21
 DB 539 HSLILEAMSHNDAYNED 556

RESULT 15
 Q8KZL5
 ID Q8KZL5 PRELIMINARY; PRT; 1554 AA.
 AC Q8KZL5;
 DT 01-OCT-2002 (TrEMBLrel. 22, Created)
 DT 01-OCT-2002 (TrEMBLrel. 22, last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, last annotation update)
 DE Glucosyltransferase.
 GN GTFU.
 OS Streptococcus sobrinus.
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 OX NCBI_TaxID=1310;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=2196684; PubMed=11960691;
 RA Hanada N., Fukushima K., Nomura Y., Sempuku H., Hayakawa M.,
 RA Mukasa H., Shiroza T., Abiko Y.;
 RT "Cloning and nucleotide sequence analysis of the Streptococcus
 RT sobrinus gtfu gene that produces a highly branched water-soluble
 RT glucan.";
 RT Biochim. Biophys. Acta 1570:75-79 (2002).
 RL EMBL; AB089438; BAC07265.1;
 DR InterPro; IPR002479; CW_binding.
 DR InterPro; IPR003318; Glyco_hydro_70.
 DR Pfam; PF01473; CW_binding_1; 6.
 DR Pfam; PF02324; Glyco_hydro_70; 1.
 KW Transferase.

SQ SEQUENCE 1554 AA; 171676 MM; 6981BCCIDAE24A73 CRC64;

Query Match 52.9%; Score 63; DB 2; Length 1554;
 Best Local Similarity 66.7%; Pred. No. 0.69;
 Matches 12; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

OY 4 HSLILEAMSDNDTPYLHD 21
 DB 488 HSLILEAMSLNDQYNED 505

Search completed: November 13, 2003, 09:43:58
 Job time : 34.6493 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 13, 2003, 09:31:40 / Search time 11.7962 Seconds
(without alignments)
154.898 Million cell updates/sec

Title: US-09-290-049a-2

Sequence: 1 VPSYSPFRAHDESEVDLIA 19

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: PIR 76:***
1: p1r1:***
2: p1r2:***
3: p1r3:***
4: p1r4:***

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	95	100.0	1475	2	B31335	gtfB protein precursor
2	91	95.8	1375	2	UT0345	dextranucrase (EC
3	84	88.4	1592	2	A38175	glucosyltransferase
4	80	84.2	1508	2	T31098	probable dextran
5	79	83.2	1290	2	UC5473	dextranucrase (EC
6	73	76.8	1365	2	A41483	glucosyltransferase
7	65	68.4	1431	2	A45866	dextranucrase (EC
8	65	68.4	1449	2	T30857	glucosyltransferase
9	65	68.4	1449	2	T30552	glucosyltransferase
10	65	68.4	1518	2	A44811	glucosyltransferase
11	64	67.4	1599	2	S22737	glucosyltransferase
12	61	64.2	1577	2	T30858	glucosyltransferase
13	46	48.4	575	2	S46329	intermediate filament
14	46	48.4	631	2	H30539	conserved hypobol
15	44	46.3	327	2	D86474	probable RING zinc
16	44	46.3	51	2	G82455	hypothetical prote
17	43	45.3	214	2	RGBSCA	regulatory protein
18	43	45.3	1260	2	A86323	protein F14D16.3
19	42	44.2	431	2	S52583	ctx protein - Erw
20	42	44.2	563	2	P81888	probable D-lactate
21	42	44.2	644	2	D81134	exoribonuclease II
22	42	44.2	644	2	A64877	RNase II, mRNA deg
23	42	44.2	644	2	C90861	RNase II, mRNA deg
24	42	44.2	644	2	B85758	ribonuclease T2 ho
25	42	44.2	659	2	A64139	probable histone t
26	42	44.2	932	2	T40216	lysine decarboxyla
27	41.5	43.7	713	2	B64743	lysine decarboxyla
28	41.5	43.7	713	2	D85503	lysine decarboxyla
29	41.5	43.7	713	2	D90652	lysine decarboxyla

30	41	43.2	145	2	G69098	hypothetical prote
31	41	43.2	344	2	JH0511	inositol 2-dehydro
32	41	43.2	597	2	F82935	DNA polymerase III
33	41	43.2	644	2	AB0656	exoribonuclease II
34	41	43.2	934	2	AG1889	WD-40 repeat prote
35	41	43.2	1146	2	E70204	hypothetical prote
36	40.5	42.6	414	2	T06120	hypothetical prote
37	40	42.1	359	2	G82197	RsfA/RsfA2 protei
38	40	42.1	431	2	B37802	crk protein - Erw
39	40	42.1	436	2	H97270	thiamine biosynthe
40	40	42.1	534	2	S41735	cholesterol estera
41	40	42.1	548	2	S32615	triacylglycerol 11
42	40	42.1	549	2	JN0552	triacylglycerol 11
43	40	42.1	549	2	JN0551	triacylglycerol 11
44	40	42.1	570	2	A90015	PTS system, lactos
45	40	42.1	572	2	B28474	phosphotransferase

ALIGNMENTS

RESULT 1
B31335
gtfB protein precursor - Streptococcus mutans
C/Species: Streptococcus mutans
C/Date: 23-Oct-1990 #sequence_revision 23-Oct-1990 #text_change 15-Oct-1999
C/Accession: B31335; A33128
R/Shiroza, T.; Ueda, S.; Kuramitsu, H.K.
U. Bacteriol. 169, 4263-4270, 1987
A/Title: Sequence analysis of the gtfB gene from Streptococcus mutans.
A/Reference number: A33135; MUID:87308013; PMID:3040685
A/Status: Preliminary
A/Accession: B31335
A/Molecule type: DNA
A/Residues: 1-1475 <SHT>
A/Cross-references: GB:M17361; NID:G153639; PIDN:AA88588.1; PID:G153640
R/Shiroza, T.; Ueda, S.; Kuramitsu, H.K.
Submitted to the Protein Sequence Database, September 1990
A/Reference number: A33128
A/Accession: A33128
A/Status: preliminary; not compared with conceptual translation
A/Molecule type: DNA
A/Residues: 1-171,173-641,'N',643-1475 <SH2>
A/Experimental source: strain GS-5
C/Superfamily: cpl repeat homology
F:1096-1115/Domain: cpl repeat homology <CP1>
F:1224-1243/Domain: cpl repeat homology <CP2>
F:1289-1308/Domain: cpl repeat homology <CP3>
F:1354-1373/Domain: cpl repeat homology <CP4>
F:1419-1438/Domain: cpl repeat homology <CP5>
Query Match 100.0%; Score 95; DB 2; Length 1475;
Best Local Similarity 100.0%; Pred. No. 1.3e-07;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VPSYSPFRAHDESEVDLIA 19
DB 552 VPSYSPFRAHDESEVDLIA 570

RESULT 2
UT0345
dextranucrase (EC 2.4.1.5) precursor - Streptococcus mutans (strain GS-5)
N/Alternate names: sucrose 6-glucosyltransferase
C/Species: Streptococcus mutans
C/Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 24-Sep-1999
C/Accession: UT0345; C33135
R/Ueda, S.; Shiroza, T.; Kuramitsu, H.K.
Gene 69, 101-109, 1988
A/Title: Sequence analysis of the gtfC gene from Streptococcus mutans GS-5.
A/Reference number: UT0345; MUID:89137980; PMID:2976010
A/Accession: UT0345
A/Molecule type: DNA

A;Residues: 1-1375 <UED>
 A;Experimental source: GS-5
 R;Shimura, T.; Ueda, S.; Kuramitsu, H.K.
 J. Bacteriol. 169, 4263-4270, 1987
 A;Title: Sequence analysis of the *gtfB* gene from *Streptococcus mutans*.
 A;Reference number: A33135; MUID:87308013; PMID:3040685
 A;Accession: C33135
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-349 <SH1>
 A;Cross-references: GB:M17361
 C;Genetics:
 A;Gene: *gtfC*
 C;Function: catalyzes the synthesis of both water-soluble and water-insoluble glucans
 A;Description: cpl repeat homology
 C;Superfamily: cpl repeat homology
 C;Keywords: duplication; glycosyltransferase; hexosyltransferase
 F;1-34/Domain: signal sequence #status predicted <SIG>
 F;135-1375/Product: glycosyltransferase #status predicted <MAT>
 F;1126-1145/Domain: cpl repeat homology <CP1>
 F;1253-1272/Domain: cpl repeat homology <CP2>
 F;1318-1337/Domain: cpl repeat homology <CP3>

Query Match 95.8%; Score 91; DB 2; Length 1375;
 Best Local Similarity 100.0%; Pred. No. 6e-07; Indels 0; Gaps 0;
 Matches 18; Conservative 0; Mismatches 0;

OY 1 VPYSFIRAHDSVODLI 18
 DB 578 VPYSFIRAHDSVODLI 595

RESULT 3
 A38175
 glucosyltransferase precursor - *Streptococcus sobrinus*
 C;Species: *Streptococcus sobrinus*
 C;Date: 28-Aug-1992 #sequence_revision 28-Aug-1992 #text_change 15-Oct-1999
 C;Accession: A38175
 R;Abu, H.; Mtsamura, T.; Kodama, T.; Ohta, H.; Fukui, K.; Kato, K.; Kagawa, H.
 J. Bacteriol. 173, 989-996, 1991
 A;Title: Peptide sequences for sucrose splitting and glucan binding within *Streptococcus*
 A;Reference number: A38175; MUID:91123227; PMID:1704006
 A;Accession: A38175
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-1592 <ABO>
 A;Cross-references: GB:D90213; NID:g217032; PIDN:BA14241.1; PID:d1014946; PID:g217033
 C;Superfamily: cpl repeat homology
 F;1093-1112/Domain: cpl repeat homology <CP1>
 F;1223-1241/Domain: cpl repeat homology <CP2>
 F;1287-1306/Domain: cpl repeat homology <CP3>
 F;1330-1351/Domain: cpl repeat homology <CP4>
 F;1352-1371/Domain: cpl repeat homology <CP5>
 F;1402-1420/Domain: cpl repeat homology <CP6>
 F;1465-1484/Domain: cpl repeat homology <CP7>
 F;1513-1532/Domain: cpl repeat homology <CP8>

Query Match 88.4%; Score 84; DB 2; Length 1592;
 Best Local Similarity 88.9%; Pred. No. 1.1e-05; Indels 0; Gaps 0;
 Matches 16; Conservative 1; Mismatches 1;

OY 1 VPYSFIRAHDSVODLI 18
 DB 548 VPYSFIRAHDSVODLI 565

RESULT 4
 T31098
 Probable dextranucrase (EC 2.4.1.5), extracellular - *Leuconostoc mesenteroides*
 C;Species: *Leuconostoc mesenteroides*
 C;Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 11-May-2000
 C;Accession: T31098
 R;Monchois, V.; Remaud-Simeon, M.; Monsan, P.; Willemot, R.M.

FEWS Microbiol. Lett. 159, 307-315, 1998
 A;Title: Cloning and sequencing of a gene coding for an extracellular dextranucrase (DS
 A;Reference number: Z20981; MUID:98164374; PMID:9503626
 A;Accession: T31098
 A;Status: preliminary; translated from GB/EMBL/DDBL
 A;Molecule type: DNA
 A;Residues: 1-1508 <MON>
 A;Cross-references: EMBL:AR030129; NID:g2766611; PID:g2766612; PIDN:AA095453.1
 A;Experimental source: strain NRRL B-1299
 C;Genetics:
 A;Gene: *dsrB*
 C;Function: produces dextran composed only of alpha(1-6) glucosidic bonds
 A;Description: produces dextran composed only of alpha(1-6) glucosidic bonds
 C;Keywords: glycosyltransferase; hexosyltransferase

Query Match 84.2%; Score 80; DB 2; Length 1508;
 Best Local Similarity 73.7%; Pred. No. 5e-05; Indels 0; Gaps 0;
 Matches 14; Conservative 4; Mismatches 1;

OY 1 VPYSFIRAHDSVODLI 19
 DB 634 IPNYSFVRAHDSVQTVIA 652

RESULT 5
 JC5473
 dextranucrase (EC 2.4.1.5) - *Leuconostoc mesenteroides*
 C;Species: *Leuconostoc mesenteroides*
 C;Date: 07-Jul-1997 #sequence_revision 29-Aug-1997 #text_change 29-Aug-1997
 C;Accession: JC5473
 R;Monchois, V.; Willemot, R.M.; Remaud-Simeon, M.; Croux, C.; Monsan, P.
 Gene 182, 23-32, 1996
 A;Title: Cloning and sequencing of a gene coding for a novel dextranucrase from *Leucon*
 A;Reference number: JC5473; MUID:97136686; PMID:8982063
 A;Accession: JC5473
 A;Status: nucleic acid sequence not shown; translation not shown
 A;Molecule type: DNA
 A;Residues: 1-1290 <MON>
 A;Cross-references: GB:U38181
 C;Comment: This enzyme catalyzes the transfer of D-glucopyranosyl units from sucrose on
 C;Genetics:
 A;Gene: *dsrA*
 C;Keywords: glycosyltransferase; hexosyltransferase
 F;78-870/Domain: catalytic #status predicted <CAT>
 F;922-1290/Domain: glucan-binding #status predicted <GCB>

Query Match 83.2%; Score 79; DB 2; Length 1290;
 Best Local Similarity 83.3%; Pred. No. 6.2e-05; Indels 0; Gaps 0;
 Matches 15; Conservative 2; Mismatches 1;

OY 2 PYSFIRAHDSVODLI 19
 DB 388 PYSFIRAHDSVODLI 405

RESULT 6
 A41483
 glucosyltransferase (EC 2.4.1.-) *gtfs* precursor - *Streptococcus sobrinus*
 C;Species: *Streptococcus sobrinus*
 C;Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 15-Oct-1999
 C;Accession: A41483
 R;Gilmore, K.S.; Russell, R.R.B.; Ferretti, J.J.
 Infect. Immun. 58, 2452-2458, 1990
 A;Title: Analysis of the *Streptococcus downii* *gtfs* gene, which specifies a glucosyltra
 A;Reference number: A41483; MUID:90316665; PMID:2142479
 A;Accession: A41483
 A;Molecule type: DNA
 A;Residues: 1-1365 <GIL>
 A;Cross-references: GB:M30943; NID:g153652; PIDN:AAA6898.1; PID:g153653
 C;Genetics:
 A;Gene: *gtfs*
 C;Superfamily: cpl repeat homology
 C;Keywords: glycosyltransferase; hexosyltransferase

Query Match 76.8%; Score 73; DB 2; Length 1365;
 Best Local Similarity 78.9%; Pred. No. 0.00071;
 Matches 15; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 VPSYFIRAHDSVQDLIA 19
 ||:|||||||
 DB 537 VPMYFIRAHDSVQTRIA 555

RESULT 7

dextranucrase (EC 2.4.1.5) precursor - Streptococcus mutans
 C/Species: Streptococcus mutans
 C/Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 07-May-1999
 C/Accession: A45866
 R/Honda, O.; Kato, C.; Kuramitsu, H.K.
 U. Gen. Microbiol. 136, 2089-2105, 1990
 A/Title: Nucleotide sequence of the Streptococcus mutans gtfD gene encoding the glucosyl
 A/Reference number: A45866; MVID:91100958; PMID:2148600
 A/Accession: A45866
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-1431 <HON>
 A/Cross-references: GB:M29296
 C/Superfamily: cpl repeat homology
 C/Keywords: glycosyltransferase; hexosyltransferase
 F:181-201/Domain: cpl repeat homology <CP1>
 F:1127-1146/Domain: cpl repeat homology <CP2>
 F:1192-1211/Domain: cpl repeat homology <CP3>
 F:1257-1276/Domain: cpl repeat homology <CP4>
 F:1277-1297/Domain: cpl repeat homology <CP5>
 F:1321-1340/Domain: cpl repeat homology <CP6>
 F:1341-1361/Domain: cpl repeat homology <CP6>
 F:1385-1406/Domain: cpl repeat homology <CP7>

Query Match 68.4%; Score 65; DB 2; Length 1431;
 Best Local Similarity 76.5%; Pred. No. 0.017;
 Matches 13; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 3 SYSFIRAHDSVQDLIA 19
 :|||
 DB 576 NYFIRAHDSVQTVIA 592

RESULT 8

glucosyltransferase - Streptococcus salivarius
 C/Species: Streptococcus salivarius
 C/Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 22-Oct-1999
 C/Accession: T30857
 R/Simpson, C.L.; Giffard, P.M.; Jacques, N.A.
 Infect. Immun. 63, 609-621, 1995
 A/Title: Streptococcus salivarius ATCC 25975 possesses at least two genes coding for pri
 A/Reference number: Z20909; MVID:95122197; PMID:7822030
 A/Accession: T30857
 A/Status: preliminary; translated from GB/EMBL/DBJ
 A/Molecule type: DNA
 A/Residues: 1-1449 <SIM>
 A/Cross-references: EMBL:L35495; NID:g662378; PID:g662379; PIDN:AAQ41412.1
 C/Genetics:
 A/Gene: gtfL

Query Match 68.4%; Score 65; DB 2; Length 1449;
 Best Local Similarity 68.8%; Pred. No. 0.018;
 Matches 11; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 3 SYSFIRAHDSVQDLIA 18
 :|||
 DB 609 NYAFVRAHDSVQSI 624

RESULT 9

T30552
 glucosyltransferase N - Streptococcus salivarius (fragment)
 C/Species: Streptococcus salivarius
 C/Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 22-Oct-1999
 C/Accession: T30552
 R/Jaffe, R.I.
 submitted to the EMBL Data Library, February 1998
 A/Description: Streptococcus salivarius V1477 gtfN.
 A/Reference number: Z20854
 A/Accession: T30552
 A/Status: preliminary; translated from GB/EMBL/DBJ
 A/Molecule type: DNA
 A/Residues: 1-1449 <JAF>
 A/Cross-references: EMBL:AF049609; NID:g293545; PID:g293546; PIDN:AAQ05156.1
 C/Genetics:
 A/Gene: gtfN

Query Match 68.4%; Score 65; DB 2; Length 1449;
 Best Local Similarity 68.8%; Pred. No. 0.018;
 Matches 11; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 3 SYSFIRAHDSVQDLIA 18
 :|||
 DB 609 NYAFVRAHDSVQSI 624

RESULT 10

glucosyltransferase (EC 2.4.1.-) I - Streptococcus salivarius
 C/Species: Streptococcus salivarius
 C/Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 15-Oct-1999
 C/Accession: A44811; S22726; S28809
 R/Giffard, P.M.; Simpson, C.L.; Milward, C.P.; Jacques, N.A.
 U. Gen. Microbiol. 137, 2577-2593, 1991
 A/Title: Molecular characterization of a cluster of at least two glucosyltransferase gen
 A/Reference number: A44811; MVID:92148377; PMID:1858391
 A/Accession: A44811
 A/Molecule type: DNA
 A/Residues: 1-1518 <GIF>
 A/Cross-references: EMBL:Z11873; NID:g47526; PIDN:CAAT7900.1; PID:g47527
 A/Note: sequence extracted from NCBI backbone (NCBI:81050, NCBI:81052)
 C/Genetics:
 A/Gene: gtfY
 C/Superfamily: cpl repeat homology
 C/Keywords: glycosyltransferase; hexosyltransferase
 F:1307-1326/Domain: cpl repeat homology <CP4>

Query Match 68.4%; Score 65; DB 2; Length 1518;
 Best Local Similarity 70.6%; Pred. No. 0.019;
 Matches 12; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 3 SYSFIRAHDSVQDLIA 19
 :|||
 DB 604 NYFIRAHDSVQDLIA 620

RESULT 11

glucosyltransferase (EC 2.4.1.-) S - Streptococcus salivarius
 C/Species: Streptococcus salivarius
 C/Date: 16-Sep-1992 #sequence_revision 16-Sep-1992 #text_change 21-Jan-2000
 C/Accession: S22737; S28810; B44811; S22727
 R/Jacques, N. the EMBL Data Library, March 1992
 submitted to the EMBL Data Library, March 1992
 A/Reference number: S22737
 A/Accession: S22737
 A/Molecule type: DNA
 A/Residues: 1-1599 <JAC>
 A/Cross-references: EMBL:Z11872; NID:g47530; PIDN:CAAT7898.1; PID:g47531
 A/Experimental source: ATCC 25975
 R/Giffard, P.M.; Simpson, C.L.; Milward, C.P.; Jacques, N.A.
 U. Gen. Microbiol. 137, 2577-2593, 1991
 A/Title: Molecular characterization of a cluster of at least two glucosyltransferase gen

PI Nichols SE;
 XX WPI: 2002-414332/44.
 DR N-PSDB; ABK52938.
 XX
 PT Glucosyltransferase B or D protein useful for producing a glucan useful
 as substitutes for and additions to modified starch and latexes in
 PT paper manufacture, comprises mutations in specific positions -
 XX
 PS Disclosure; Page 21-25; 44pp; English.

XX The invention an isolated protein comprising a glucosyltransferase
 CC (GTF) B polypeptide having changes at position from 1448V, D457N,
 CC D567T, K1014T, D457N/D567T, D457N/D571K, D567T/D571K,
 CC D567T/D571K/K1014T, 1448V/D457N/D567T/D571K/K779Q/K1014T,
 CC Y169A/Y170A/Y171A, and K779Q or a GTF D polypeptide having
 CC changes at positions from T589D, T589E, N471D, N471D/T589D, and
 CC N471D/T589E. Also included are a glucan produced by the GTF mutant,
 CC an isolated polynucleotide which encodes P1 or P2, or its complementary
 CC polynucleotide, a ribonucleic acid sequence encoding the GTF mutant,
 CC an expression cassette comprising the polynucleotide operably linked to a
 CC promoter, a vector comprising the expression cassette, host cell
 CC introduced with the vector, a transgenic plant comprising the
 CC vector, a seed or tuber from the transgenic plant, a paper sizing and/or
 CC coating composition comprising a glucan produced in a plant transformed
 CC with a gene encoding the mutant GTF, wild type or, starch, a latex,
 CC thermoplastic molecule or their combinations or glucan and starch where
 CC the glucan is produced in the amyloplast and/or vacuole or a maize line
 CC deficient in starch biosynthesis, transformed with a gene encoding a
 CC glucosyltransferase B or D enzyme, wild-type or mutant and a paper
 CC comprising the glucan (paper sizing/coating agent). The vector is useful
 CC for producing a glucan in a plant. The method comprises transforming a
 CC plant cell with the vector, growing the plant cell under plant growing
 CC conditions to produce a regenerated plant and inducing expression of the
 CC polynucleotide for a time sufficient to produce the glucan in the
 CC regenerated plant, where the vector contains a transit sequence from
 CC ribulose biphosphate carboxylase small subunit, waxy, brittle-1 and
 CC chlorophyll AB binding protein to produce a transgenic plant, and glucan
 CC is produced in the amyloplast of potato or the vacuole of sugar beet.
 CC Glucans are useful as substitutes for and additions to modified starch
 CC and latexes in paper manufacture. Unlike prior art techniques, which
 CC require input materials that produce chemical effluents, paper
 CC manufacture utilizing the glucan produced by GTF, which utilizes
 CC biologically produced input materials, is more cost-effective and
 CC environmentally friendly. Moreover, glucans also exhibit thermoplastic
 CC properties and impart gloss to the paper during coating step.
 CC The present sequence represents GTFB.

XX Sequence 1475 AA;

Query Match 100.0%; Score 95; DB 23; Length 1475;
 Best Local Similarity 100.0%; Pred. No. 2.9e-08;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VPSYSTRAHDSVQDLIA 19
 DB 552 VPSYSTRAHDSVQDLIA 570

RESULT 2
 AAU98030
 ID AAU98030 standard; Protein; 1475 AA.

XX AAU98030;

DT 27-AUG-2002 (first entry)

DE S. mutans glucosyltransferase GTFB mutant I448V.

XX Glucosyltransferase; GTFB; transgenic plant; paper sizing;
 KW coating composition; glucan; starch; latex; thermoplastic molecule;
 KM amyloplast; vacuole; paper manufacture; mutant; mutin.
 XX

OS Streptococcus mutans.
 OS Synthetic.
 XX Key Location/Qualifiers
 FH MISC-difference 448 /note= "Wild-type Ile substituted by Val"
 FT US2002031826-A1.
 XX
 XX 14-MAR-2002.
 PD
 PD 19-DEC-2000; 2000US-0740274.
 PF
 XX 11-DEC-1998; 98US-0210361.
 XX 07-JUN-1995; 95US-0478704.
 PR 07-JUN-1995; 95US-0482711.
 PR 07-JUN-1995; 95US-0485243.
 PR 16-JAN-1998; 98US-0007999.
 PR 16-JAN-1998; 98US-0008172.
 PR 20-JAN-1998; 98US-0009620.
 XX
 XX (NICH/) NICHOLS S E.
 PA
 PI Nichols SE;
 XX WPI: 2002-414332/44.
 DR
 XX Glucosyltransferase B or D protein useful for producing a glucan useful
 PT as substitutes for and additions to modified starch and latexes in
 PT paper manufacture, comprises mutations in specific positions -
 XX
 PS Claim 36; Page -; 44pp; English.

XX The invention an isolated protein comprising a glucosyltransferase
 CC (GTF) B polypeptide having changes at position from 1448V, D457N,
 CC D567T, K1014T, D457N/D567T, D457N/D571K, D567T/D571K,
 CC D567T/D571K/K1014T, 1448V/D457N/D567T/D571K/K779Q/K1014T,
 CC Y169A/Y170A/Y171A, and K779Q or a GTF D polypeptide having
 CC changes at positions from T589D, T589E, N471D, N471D/T589D, and
 CC N471D/T589E. Also included are a glucan produced by the GTF mutant,
 CC an isolated polynucleotide which encodes P1 or P2, or its complementary
 CC polynucleotide, a ribonucleic acid sequence encoding the GTF mutant,
 CC an expression cassette comprising the polynucleotide operably linked to a
 CC promoter, a vector comprising the expression cassette, host cell
 CC introduced with the vector, a transgenic plant comprising the
 CC vector, a seed or tuber from the transgenic plant, a paper sizing and/or
 CC coating composition comprising a glucan produced in a plant transformed
 CC with a gene encoding the mutant GTF, wild type or, starch, a latex,
 CC thermoplastic molecule or their combinations or glucan and starch where
 CC the glucan is produced in the amyloplast and/or vacuole or a maize line
 CC deficient in starch biosynthesis, transformed with a gene encoding a
 CC glucosyltransferase B or D enzyme, wild-type or mutant and a paper
 CC comprising the glucan (paper sizing/coating agent). The vector is useful
 CC for producing a glucan in a plant. The method comprises transforming a
 CC plant cell with the vector, growing the plant cell under plant growing
 CC conditions to produce a regenerated plant and inducing expression of the
 CC polynucleotide for a time sufficient to produce the glucan in the
 CC regenerated plant, where the vector contains a transit sequence from
 CC ribulose biphosphate carboxylase small subunit, waxy, brittle-1 and
 CC chlorophyll AB binding protein to produce a transgenic plant, and glucan
 CC is produced in the amyloplast of potato or the vacuole of sugar beet.
 CC Glucans are useful as substitutes for and additions to modified starch
 CC and latexes in paper manufacture. Unlike prior art techniques, which
 CC require input materials that produce chemical effluents, paper
 CC manufacture utilizing the glucan produced by GTF, which utilizes
 CC biologically produced input materials, is more cost-effective and
 CC environmentally friendly. Moreover, glucans also exhibit thermoplastic
 CC properties and impart gloss to the paper during coating step.
 CC The present sequence represents a GTFB mutant of the invention.
 CC Note: The present sequence is not shown in the specification but
 CC was created by the indexer using the GTFB sequence appearing as AAU98027
 CC and the information in claim 36.
 XX

SQ Sequence 1475 AA;

Query Match 100.0%; Score 95; DB 23; Length 1475;
 Best Local Similarity 100.0%; Pred. No. 2.9e-08;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VPSYSFIRAHSEVQDLIA 19
 |||||
 DB 552 VPSYSFIRAHSEVQDLIA 570

RESULT 3
 AA098031
 ID AAU98031 standard; Protein; 1475 AA.
 XX AAU98031;
 AC
 XX
 XX 27-AUG-2002 (first entry)
 DT
 XX
 XX S. mutans glucosyltransferase GTFB mutant D457N.
 DE
 XX Glucosyltransferase; GTFB; transgenic plant; paper sizing;
 KW coating composition; glucan; starch; latex; thermoplastic molecule;
 KM amyloplast; vacuole; paper manufacture; mutant; mutein.
 XX
 XX Streptococcus mutans.
 OS Synthetic.
 OS
 XX
 XX Key Location/Qualifiers
 FH Misc-difference 457 /note="Wild-type Asp substituted by Asn"
 FT
 XX
 XX US2002031826-A1.
 PN
 XX
 PD 14-MAR-2002.
 XX
 PF 19-DEC-2000; 2000US-0740274.
 XX
 XX 11-DEC-1998; 98US-0210361.
 PR 07-JUN-1995; 95US-0478704.
 PR 07-JUN-1995; 95US-0482711.
 PR 07-JUN-1995; 95US-0485243.
 PR 16-JAN-1998; 98US-0007999.
 PR 16-JAN-1998; 98US-0008172.
 PR 20-JAN-1998; 98US-0009620.
 XX
 XX (NICH/) NICHOLS S E.
 PA
 XX Nichols SE;
 PI
 XX
 DR WPI; 2002-414332/44.
 XX
 XX Glucosyltransferase B or D protein useful for producing a glucan useful
 PT as substitutes for and additions to modified starch and latexes in
 PT paper manufacture, comprises mutations in specific positions
 XX
 XX Claim 36; Page -; 44p; English.
 PS
 XX
 CC The invention an isolated protein comprising a glucosyltransferase
 CC (GTF) B polypeptide having changes at position from I448V, D457N,
 CC D567T, K1014T, D457N/D567T, D457N/D571K, D567T/D571K,
 CC D567T/D571K/K1014T, I448V/D457N/D567T/D571K/K779Q/K1014T,
 CC Y169A/Y170A/Y171A, and K779Q or a GTF D polypeptide having
 CC changes at positions from T589D, T589E, N471D, N471D/T589D, and
 CC N471D/T589E. Also included are a glucan produced by the GTF mutant,
 CC an isolated polynucleotide which encodes PI or P2, or its complementary
 CC polynucleotide, a ribonucleic acid sequence encoding the GTF mutant,
 CC an expression cassette comprising the polynucleotide operably linked to a
 CC promoter, a vector comprising the expression cassette, host cell
 CC introduced with the vector, a transgenic plant comprising the
 CC vector, a seed or tuber from the transgenic plant, a paper sizing and/or
 CC coating composition comprising a glucan produced in a plant transformed
 CC with a gene encoding the mutant GTF, wild type or, starch, a latex.

CC thermoplastic molecule or their combinations or glucan and starch where
 CC the glucan is produced in the amyloplast and/or vacuole or a maize line
 CC deficient in starch biosynthesis, transformed with a gene encoding a
 CC glucosyltransferase B or D enzyme, wild-type or mutant and a paper
 CC comprising the glucan (paper sizing/coating agent). The vector is useful
 CC for producing a glucan in a plant. The method comprises transforming a
 CC plant cell with the vector, growing the plant cell under plant growing
 CC conditions to produce a regenerated plant and inducing expression of the
 CC polynucleotide for a time sufficient to produce the glucan in the
 CC regenerated plant, where the vector contains a transit sequence from
 CC ribulose biphosphate carboxylase small subunit, waxy, brittle-1 and
 CC chlorophyll a/b binding protein to produce a transgenic plant, and glucan
 CC is produced in the amyloplast of potato or the vacuole of sugar beet.
 CC Glucans are useful as substitutes for and additions to modified starch
 CC and latexes in paper manufacture. Unlike prior art techniques, which
 CC require input materials that produce chemical effluents, paper
 CC manufacture utilizing the glucan produced by GTF, which utilizes
 CC biologically produced input materials, is more cost-effective and
 CC environmentally friendly. Moreover, glucans also exhibit thermoplastic
 CC properties and impart gloss to the paper during coating step.
 CC The present sequence represents a GTFB mutant of the invention.
 CC Note: The present sequence is not shown in the specification but
 CC was created by the indexer using the GTFB sequence appearing as AAU98027
 CC and the information in claim 36.

SQ Sequence 1475 AA;

Query Match 100.0%; Score 95; DB 23; Length 1475;
 Best Local Similarity 100.0%; Pred. No. 2.9e-08;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VPSYSFIRAHSEVQDLIA 19
 |||||
 DB 552 VPSYSFIRAHSEVQDLIA 570

RESULT 4
 AAU98033
 ID AAU98033 standard; Protein; 1475 AA.
 XX AAU98033;
 AC
 XX
 XX 27-AUG-2002 (first entry)
 DT
 XX
 XX S. mutans glucosyltransferase GTFB mutant K1014T.
 DE
 XX Glucosyltransferase; GTFB; transgenic plant; paper sizing;
 KW coating composition; glucan; starch; latex; thermoplastic molecule;
 KM amyloplast; vacuole; paper manufacture; mutant; mutein.
 XX
 XX Streptococcus mutans.
 OS Synthetic.
 OS
 XX
 XX Key Location/Qualifiers
 FH Misc-difference 1014 /note="Wild-type Lys substituted by Thr"
 FT
 XX
 XX US2002031826-A1.
 PN
 XX
 PD 14-MAR-2002.
 XX
 PF 19-DEC-2000; 2000US-0740274.
 XX
 XX 11-DEC-1998; 98US-0210361.
 PR 07-JUN-1995; 95US-0478704.
 PR 07-JUN-1995; 95US-0482711.
 PR 07-JUN-1995; 95US-0485243.
 PR 16-JAN-1998; 98US-0007999.
 PR 16-JAN-1998; 98US-0008172.
 PR 20-JAN-1998; 98US-0009620.
 XX
 XX (NICH/) NICHOLS S E.
 PA
 XX Nichols SE;
 PI
 XX
 DR WPI; 2002-414332/44.
 XX
 XX Glucosyltransferase B or D protein useful for producing a glucan useful
 PT as substitutes for and additions to modified starch and latexes in
 PT paper manufacture, comprises mutations in specific positions
 XX
 XX Claim 36; Page -; 44p; English.
 PS
 XX
 CC The invention an isolated protein comprising a glucosyltransferase
 CC (GTF) B polypeptide having changes at position from I448V, D457N,
 CC D567T, K1014T, D457N/D567T, D457N/D571K, D567T/D571K,
 CC D567T/D571K/K1014T, I448V/D457N/D567T/D571K/K779Q/K1014T,
 CC Y169A/Y170A/Y171A, and K779Q or a GTF D polypeptide having
 CC changes at positions from T589D, T589E, N471D, N471D/T589D, and
 CC N471D/T589E. Also included are a glucan produced by the GTF mutant,
 CC an isolated polynucleotide which encodes PI or P2, or its complementary
 CC polynucleotide, a ribonucleic acid sequence encoding the GTF mutant,
 CC an expression cassette comprising the polynucleotide operably linked to a
 CC promoter, a vector comprising the expression cassette, host cell
 CC introduced with the vector, a transgenic plant comprising the
 CC vector, a seed or tuber from the transgenic plant, a paper sizing and/or
 CC coating composition comprising a glucan produced in a plant transformed
 CC with a gene encoding the mutant GTF, wild type or, starch, a latex.

PI Nichols SE;
 XX
 DR WPI; 2002-414332/44.
 XX
 PT Glucosyltransferase B or D protein useful for producing a glucan useful
 PT as substitutes for and additions to modified starch and latexes in
 PT paper manufacture, comprises mutations in specific positions -
 XX
 PS Claim 36; Page -; 44pp; English.
 XX
 CC The invention an isolated protein comprising a glucosyltransferase
 CC (GTF) B polypeptide having changes at position from 1448V, D457N,
 CC D567T, K1014T, D457N/D567T, D457N/D571K, D567T/D571K,
 CC D567T/D571K/K1014T, 1448V/D457N/D567T/D571K/K719Q/K1014T,
 CC Y169A/Y170A/Y171A, and K719Q or a GTF D polypeptide having
 CC changes at positions from T589D, T589E, N471D, N471D/T589D, and
 CC N471D/T589E. Also included are a glucan produced by the GTF mutant,
 CC an isolated polynucleotide which encodes P1 or P2, or its complementary
 CC polynucleotide, a ribonucleic acid sequence encoding the GTF mutant,
 CC an expression cassette comprising the polynucleotide operably linked to a
 CC promoter, a vector comprising the expression cassette, host cell
 CC introduced with the vector, a transgenic plant comprising the
 CC vector, a seed or tuber from the transgenic plant, a paper sizing and/or
 CC coating composition comprising a glucan produced in a plant transformed
 CC with a gene encoding the mutant GTF, wild type or, starch, a latex,
 CC thermoplastic molecule or their combinations or glucan and starch where
 CC the glucan is produced in the amyloplast and/or vacuole or a maize line
 CC deficient in starch biosynthesis, transformed with a gene encoding a
 CC glucosyltransferase B or D enzyme, wild-type or mutant and a paper
 CC comprising the glucan (paper sizing/coating agent). The vector is useful
 CC for producing a glucan in a plant. The method comprises transforming a
 CC plant cell with the vector, growing the plant cell under plant growing
 CC conditions to produce a regenerated plant and inducing expression of the
 CC polynucleotide for a time sufficient to produce the glucan in the
 CC regenerated plant, where the vector contains a transit sequence from
 CC ribulose biphosphate carboxylase small subunit, waxy, brittle-1 and
 CC chlorophyll AB binding protein to produce a transgenic plant, and glucan
 CC is produced in the amyloplast of potato or the vacuole of sugar beet.
 CC Glucans are useful as substitutes for and additions to modified starch
 CC and latexes in paper manufacture. Unlike prior art techniques, which
 CC require input materials that produce chemical effluents, paper
 CC manufacture utilizing the glucan produced by GTF, which utilizes
 CC biologically produced input materials, is more cost-effective and
 CC environmentally friendly. Moreover, glucans also exhibit thermoplastic
 CC properties and impart gloss to the paper during coating step.
 CC The present sequence represents a GTFB mutant of the invention.
 CC Note: The present sequence is not shown in the specification but
 CC was created by the indexer using the GTFB sequence appearing as AAU98027
 CC and the information in claim 36.
 XX
 SQ Sequence 1475 AA;
 XX
 QY
 Db 1 VPSYSFIRAHDSFVQDLIA 19
 552 VPSYSFIRAHDSFVQDLIA 570
 100.0%; Score 95; DB 23; Length 1475;
 Best Local Similarity 100.0%; Pred. No. 2.9e-08;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 RESULT 5
 AAU98035 standard; Protein; 1475 AA.
 XX AAU98035;
 AC AAU98035;
 DT 27-AUG-2002 (first entry)
 XX
 DE 5. mutans glucosyltransferase GTFB mutant D457N/D571K.
 XX
 KM Glucosyltransferase; GTFB; transgenic plant; paper sizing;
 KM coating composition; glucan; starch; latex; thermoplastic molecule;

KM amyloplast; vacuole; paper manufacture; mutant; mutein.
 XX
 OS Streptococcus mutans.
 OS Synthetic.
 XX
 FH Key Location/Qualifiers
 FT Misc-difference 457 /note= "Wild-type Asp substituted by Asn"
 FT Misc-difference 571 /note= "Wild-type Asp substituted by Lys"
 FT
 XX
 XX US2002031826-A1.
 XX
 XX 14-MAR-2002.
 XX
 XX 19-DEC-2000; 2000US-0740274.
 XX
 XX 11-DEC-1998; 98US-0210361.
 XX 07-JUN-1995; 95US-0478704.
 XX 07-JUN-1995; 95US-0482711.
 XX 16-JAN-1998; 98US-0007999.
 XX 16-JAN-1998; 98US-0008172.
 XX 20-JAN-1998; 98US-0009620.
 XX
 XX (NICH/) NICHOLS S E.
 XX
 PI Nichols SE;
 DR WPI; 2002-414332/44.
 XX
 PT Glucosyltransferase B or D protein useful for producing a glucan useful
 PT as substitutes for and additions to modified starch and latexes in
 PT paper manufacture, comprises mutations in specific positions -
 XX
 PS Claim 36; Page -; 44pp; English.
 XX
 CC The invention an isolated protein comprising a glucosyltransferase
 CC (GTF) B polypeptide having changes at position from 1448V, D457N,
 CC D567T, K1014T, D457N/D567T, D457N/D571K, D567T/D571K,
 CC D567T/D571K/K1014T, 1448V/D457N/D567T/D571K/K719Q/K1014T,
 CC Y169A/Y170A/Y171A, and K719Q or a GTF D polypeptide having
 CC changes at positions from T589D, T589E, N471D, N471D/T589D, and
 CC N471D/T589E. Also included are a glucan produced by the GTF mutant,
 CC an isolated polynucleotide which encodes P1 or P2, or its complementary
 CC polynucleotide, a ribonucleic acid sequence encoding the GTF mutant,
 CC an expression cassette comprising the polynucleotide operably linked to a
 CC promoter, a vector comprising the expression cassette, host cell
 CC introduced with the vector, a transgenic plant comprising the
 CC vector, a seed or tuber from the transgenic plant, a paper sizing and/or
 CC coating composition comprising a glucan produced in a plant transformed
 CC with a gene encoding the mutant GTF, wild type or, starch, a latex,
 CC thermoplastic molecule or their combinations or glucan and starch where
 CC the glucan is produced in the amyloplast and/or vacuole or a maize line
 CC deficient in starch biosynthesis, transformed with a gene encoding a
 CC glucosyltransferase B or D enzyme, wild-type or mutant and a paper
 CC comprising the glucan (paper sizing/coating agent). The vector is useful
 CC for producing a glucan in a plant. The method comprises transforming a
 CC plant cell with the vector, growing the plant cell under plant growing
 CC conditions to produce a regenerated plant and inducing expression of the
 CC polynucleotide for a time sufficient to produce the glucan in the
 CC regenerated plant, where the vector contains a transit sequence from
 CC ribulose biphosphate carboxylase small subunit, waxy, brittle-1 and
 CC chlorophyll AB binding protein to produce a transgenic plant, and glucan
 CC is produced in the amyloplast of potato or the vacuole of sugar beet.
 CC Glucans are useful as substitutes for and additions to modified starch
 CC and latexes in paper manufacture. Unlike prior art techniques, which
 CC require input materials that produce chemical effluents, paper
 CC manufacture utilizing the glucan produced by GTF, which utilizes
 CC biologically produced input materials, is more cost-effective and
 CC environmentally friendly. Moreover, glucans also exhibit thermoplastic
 CC properties and impart gloss to the paper during coating step.
 CC The present sequence represents a GTFB mutant of the invention.

CC Note: The present sequence is not shown in the specification but
CC was created by the indexer using the GTFB sequence appearing as AAU98027
CC and the information in claim 36.

XX SQ Sequence 1475 AA;

Query Match 100.0%; Score 95; DB 23; Length 1475;
Best Local Similarity 100.0%; Pred. No. 2.9e-08;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VPSYSFIRAHSEVODLIA 19
|||
552 VPSYSFIRAHSEVODLIA 570

RESULT 6
AAU98039
ID AAU98039 standard; Protein; 1475 AA.

XX AC AAU98039;

XX DT 27-AUG-2002 (first entry)

XX S. mutans glucosyltransferase GTFB mutant YYY169-171AAA.

XX Glucosyltransferase; GTFB; transgenic plant; paper sizing;
XX coating composition; glucan, starch; latex; thermoplastic molecule;
XX amyloplast; vacuole; paper manufacture; mutant; mutein.

XX Streptococcus mutans.
XX Synthetic.

XX Key Location/Qualifiers
FT Misc-difference 169..171
FT /note= "Wild-type Tyr-Tyr substituted by
FT Ala-Ala-Ala"

XX US2002031826-A1.

XX 14-MAR-2002.

XX 19-DEC-2000; 2000US-0740274.

XX 11-DEC-1998; 98US-0210361.

XX 07-JUN-1995; 95US-0478704.

XX 07-JUN-1995; 95US-0482711.

XX 16-JAN-1998; 98US-0007999.

XX 16-JAN-1998; 98US-0008172.

XX 20-JAN-1998; 98US-0009620.

XX (NICH/) NICHOLS S E.

XX Nichols SE;

XX WPI; 2002-414332/44.

XX Claim 36; Page -; 44pp; English.

XX The invention an isolated protein comprising a glucosyltransferase

XX (GTF) B polypeptide having changes at position from 1448V, D457N,
XX D567T, K1014T, D457N/D567T, D457N/D571K, D567T/D571K,

XX D567T/D571K/K1014T, 1448V/D457N/D567T/K779Q/K1014T,
XX Y169A/Y170A/Y171A, and K779Q or a GTF D polypeptide having

XX changes at positions from T589D, T589E, N471D, N471D/T589D, and
XX N471D/T589E. Also included are a glucan produced by the GTF mutant,

XX an isolated polynucleotide which encodes P1 or P2, or its complementary
XX polynucleotide, a ribonucleic acid sequence encoding the GTF mutant,
XX an expression cassette comprising the polynucleotide operably linked to a

CC promoter, a vector comprising the expression cassette, host cell
CC introduced with the vector, a transgenic plant comprising the
CC vector, a seed or tuber from the transgenic plant, a paper sizing and/or
CC coating composition comprising a glucan produced in a plant transformed
CC with a gene encoding the mutant GTF, wild type or, starch, a latex,
CC thermoplastic molecule or their combinations or glucan and starch where
CC the glucan is produced in the amyloplast and/or vacuole or a maize line
CC deficient in starch biosynthesis, transformed with a gene encoding a
CC glucosyltransferase B or D enzyme, wild-type or mutant and a paper
CC comprising the glucan (paper sizing/coating agent). The vector is useful
CC for producing a glucan in a plant. The method comprises transforming a
CC plant cell with the vector, growing the plant cell under plant growing
CC conditions to produce a regenerated plant and inducing expression of the
CC polynucleotide for a time sufficient to produce the glucan in the
CC regenerated plant, where the vector contains a transit sequence from
CC ribulose biphosphate carboxylase small subunit, waxy, brittle-1 and
CC chlorophyll AB binding protein to produce a transgenic plant, and glucan
CC is produced in the amyloplast of potato or the vacuole of sugar beet.
CC Glucans are useful as substitutes for and additions to modified starch
CC and latexes in paper manufacture. Unlike prior art techniques, which
CC require input materials that produce chemical effluents, paper
CC manufacture utilizing the glucan produced by GTF, which utilizes
CC biologically produced input materials, is more cost-effective and
CC environmentally friendly. Moreover, glucans also exhibit thermoplastic
CC properties and impart gloss to the paper during coating step.
CC The present sequence represents a GTFB mutant of the invention but
CC Note: The present sequence is not shown in the specification but
CC was created by the indexer using the GTFB sequence appearing as AAU98027
CC and the information in claim 36.

XX SQ Sequence 1475 AA;

Query Match 100.0%; Score 95; DB 23; Length 1475;
Best Local Similarity 100.0%; Pred. No. 2.9e-08;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VPSYSFIRAHSEVODLIA 19
|||
552 VPSYSFIRAHSEVODLIA 570

RESULT 7
AAU98040
ID AAU98040 standard; Protein; 1475 AA.

XX AAU98040;

XX DT 27-AUG-2002 (first entry)

XX S. mutans glucosyltransferase GTFB mutant K779Q.

XX Glucosyltransferase; GTFB; transgenic plant; paper sizing;
XX coating composition; glucan, starch; latex; thermoplastic molecule;
XX amyloplast; vacuole; paper manufacture; mutant; mutein.

XX Streptococcus mutans.
XX Synthetic.

XX Key Location/Qualifiers
FT Misc-difference 779
FT /note= "Wild-type Lys substituted by Gln"

XX US2002031826-A1.

XX 14-MAR-2002.

XX 19-DEC-2000; 2000US-0740274.

XX 11-DEC-1998; 98US-0210361.

XX 07-JUN-1995; 95US-0478704.

XX 07-JUN-1995; 95US-0482711.

XX 16-JAN-1998; 98US-0007999.

DR WPI; 2002-448101/48.
 XX Anti-carries agent composed of a monoclonal antibody against an
 PT inhibitory enzyme against water insoluble glucan synthetase of glucosyl
 PT transferase-B (GTF-B) of Streptococcus mutans -
 XX Claim 3; Page 13-16; 28pp; Japanese.
 XX The invention relates to a monoclonal antibody against dental caries and
 CC an anti-carries agent composed of a monoclonal antibody produced by
 CC Streptococcus mutans, particularly mouse-hybridoma MHP126 (FERM P-17566)
 CC or mouse-hybridoma MHP136 (FERM P-17567), against an enzyme having
 CC inhibitive activity against water insoluble glucan synthetase of glucosyl
 CC transferase-B. The monoclonal antibody specifically inhibits water
 CC insoluble glucan synthetase of Streptococcus mutans produced glucosyl
 CC transferase-B and is used in the immunotherapy of dental caries. This
 CC sequence represents a Streptococcus mutans monoclonal antibody-related
 CC protein.
 XX Sequence 1476 AA;
 SQ
 Query Match 95.8%; Score 91; DB 23; Length 1476;
 Best Local Similarity 100.0%; Pred. No. 1.6e-07;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 VPSYSFIRAHDSVODLI 18
 DB 552 VPSYSFIRAHDSVODLI 569
 RESULT 12
 AAU98032 standard; Protein; 1475 AA.
 XX AAU98032;
 AC 27-AUG-2002 (first entry)
 XX
 DT S. mutans glucosyltransferase GTFB mutant D567T.
 DB
 XX Glucosyltransferase; GTFB; transgenic plant; paper sizing;
 KW coating composition; glucan; starch; latex; thermoplastic molecule;
 KM amyloplast; vacuole; paper manufacture; mutant; mutein.
 XX
 OS Streptococcus mutans.
 OS Synthetic.
 XX
 FH Key Location/Qualifiers
 FT Misc-difference 567 /note="Wild-type Asp substituted by Thr"
 FT
 XX US2002031826-A1.
 PN
 XX 14-MAR-2002.
 PD
 XX 19-DEC-2000; 2000US-0740274.
 PF
 XX 11-DEC-1998; 98US-0210361.
 PR 07-JUN-1995; 95US-0478704.
 PR 07-JUN-1995; 95US-0482711.
 PR 07-JUN-1995; 95US-0485243.
 PR 16-JAN-1998; 98US-0007999.
 PR 16-JAN-1998; 98US-0008172.
 PR 20-JAN-1998; 98US-0009620.
 PR
 XX (NICH/) NICHOLS S E.
 PA
 PI Nichols SE;
 XX
 DR WPI; 2002-414332/44.
 XX
 PT Glucosyltransferase B or D protein useful for producing a glucan useful
 as substitutes for and additions to modified starch and latexes in

PT paper manufacture, comprises mutations in specific positions -
 XX Claim 36; Page -; 44pp; English.
 PS The invention an isolated protein comprising a glucosyltransferase
 XX (GTF) B polypeptide having changes at position from 1448V, D457N,
 CC D567T, K1014T, D457N/D567T, D457N/D571K, D567T/D571K,
 CC D567T/D571K/K1014T, D448V/D457N/D567T/K779C/K1014T,
 CC Y159A/Y170A/Y171A, and K779C or a GTF D polypeptide having
 CC changes at positions from T589D, T589E, N471D, N471D/T589D and
 CC N471D/T589E. Also included are a glucan produced by the GTF mutant,
 CC an isolated polynucleotide which encodes P1 or P2, or its complementary
 CC polynucleotide, a ribonucleic acid sequence encoding the GTF mutant,
 CC an expression cassette comprising the polynucleotide operably linked to a
 CC promoter, a vector comprising the expression cassette, host cell
 CC introduced with the vector, a transgenic plant comprising the
 CC vector, a seed or tuber from the transgenic plant, a paper transformed
 CC coating composition comprising a glucan produced in a plant transformed
 CC with a gene encoding the mutant GTF, wild type or, starch, a latex,
 CC thermoplastic molecule or their combinations and/or vacuole or a maize line
 CC the glucan is produced in the amyloplast and/or vacuole of a plant where
 CC deficient in starch biosynthesis, transformed with a gene encoding a
 CC glucosyltransferase B or D enzyme, wild-type or mutant and a paper
 CC comprising the glucan (paper sizing/coating agent). The vector is useful
 CC for producing a glucan in a plant. The method comprises transforming a
 CC plant cell with the vector, growing the plant cell under plant growing
 CC conditions to produce a regenerated plant and inducing expression of the
 CC polynucleotide for a time sufficient to produce the glucan in the
 CC regenerated plant, where the vector contains a trans sequence from
 CC ribulose biphosphate carboxylase small subunit, waxy, brittle-1 and
 CC chlorophyll AB binding protein to produce a transgenic plant, and glucan
 CC is produced in the amyloplast of potato or the vacuole of sugar beet.
 CC Glucans are useful as substitutes for and additions to modified starch
 CC and latexes in paper manufacture, unlike prior art techniques, which
 CC require input materials that produce chemical effluents, paper
 CC manufacture utilizing the glucan produced by GTF, which utilizes
 CC biologically produced input materials, is more cost-effective and
 CC environmentally friendly. Moreover, glucans also exhibit thermoplastic
 CC properties and impart gloss to the paper during coating step.
 CC The present sequence represents a GTFB mutant of the invention but
 CC Note: The present sequence is not shown in the specification but
 CC was created by the indexer using the GTFB sequence appearing as AAU98027
 CC and the information in claim 36.
 XX
 SQ Sequence 1475 AA;
 Query Match 92.6%; Score 88; DB 23; Length 1475;
 Best Local Similarity 94.7%; Pred. No. 5.8e-07;
 Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 VPSYSFIRAHDSVODLI 19
 DB 552 VPSYSFIRAHDSVODLI 570
 RESULT 13
 AAU98034 standard; Protein; 1475 AA.
 XX AAU98034;
 AC 27-AUG-2002 (first entry)
 XX
 DT S. mutans glucosyltransferase GTFB mutant D457N/D567T.
 DB
 XX Glucosyltransferase; GTFB; transgenic plant; paper sizing;
 KW coating composition; glucan; starch; latex; thermoplastic molecule;
 KM amyloplast; vacuole; paper manufacture; mutant; mutein.
 XX
 OS Streptococcus mutans.
 OS Synthetic.
 XX
 FH Key Location/Qualifiers

FT Misc-difference 457 /note= "Wild-type Asp substituted by Asn"
 FT Misc-difference 567 /note= "Wild-type Asp substituted by Thr"
 FT US2002031826-A1.
 XX 14-MAR-2002.
 XX 19-DEC-2000; 2000US-0740274.
 XX 11-DEC-1998; 98US-0210361.
 XX 07-JUN-1995; 95US-0478704.
 XX 07-JUN-1995; 95US-0482711.
 XX 07-JUN-1995; 95US-0485243.
 XX 16-JAN-1998; 98US-0007999.
 XX 16-JAN-1998; 98US-0008172.
 XX 20-JAN-1998; 98US-0009620.
 XX (NICH/) NICHOLS S E.
 XX Nichols SE;
 XX WPI; 2002-414332/44.
 DR Glucosyltransferase B or D protein useful for producing a glucan useful
 PT as substitutes for and additions to modified starch and latexes in
 PT paper manufacture, comprises mutations in specific positions -
 XX Claim 36; Page -, 44pp; English.
 XX The invention an isolated protein comprising a glucosyltransferase
 CC (GTF) B polypeptide having changes at position from 1448V, D457N,
 CC D567T, K1014T, D457N/D567T, D457N/D571K, D567T/D571K,
 CC D567T/D571K/K1014T, I448V/D457N/D571K/K719Q/K1014T,
 CC Y169A/Y170A/Y171A, and K779Q or a GTF D polypeptide having
 CC changes at positions from T589D, T589E, N471D, N471D/T589D, and
 CC N471D/T589E. Also included are a glucan produced by the GTF mutant,
 CC an isolated polynucleotide which encodes P1 or P2, or its complementary
 CC polynucleotide, a ribonucleic acid sequence encoding the GTF mutant,
 CC an expression cassette comprising the polynucleotide operably linked to a
 CC promoter, a vector comprising the expression cassette, host cell
 CC introduced with the vector, a transgenic plant comprising the
 CC vector, a seed or tuber from the transgenic plant, a paper sizing and/or
 CC coating composition comprising a glucan produced in a plant transformed
 CC with a gene encoding the mutant GTF, wild type or, starch, a latex,
 CC thermoplastic molecule or their combinations or glucan and starch where
 CC the glucan is produced in the amyloplast and/or vacuole or a maize line
 CC deficient in starch biosynthesis, transformed with a gene encoding a
 CC glucosyltransferase B or D enzyme, wild-type or mutant and a paper
 CC comprising the glucan (paper sizing/coating agent). The vector is useful
 CC for producing a glucan in a plant. The method comprises transforming a
 CC plant cell with the vector, growing the plant cell under plant growing
 CC conditions to produce a regenerated plant and inducing expression of the
 CC polynucleotide for a time sufficient to produce the glucan in the
 CC regenerated plant, where the vector contains a transit sequence from
 CC ribulose biphosphate carboxylase small subunit, waxy, brittle-1 and
 CC chlorophyll AB binding protein to produce a transgenic plant, and glucan
 CC is produced in the amyloplast of potato or the vacuole of sugar beet.
 CC Glucans are useful as substitutes for and additions to modified starch
 CC and latexes in paper manufacture. Unlike prior art techniques, which
 CC require input materials that produce chemical effluents, paper
 CC manufacture utilizing the glucan produced by GTF, which utilizes
 CC biologically produced input materials, is more cost-effective and
 CC environmentally friendly. Moreover, glucans also exhibit thermoplastic
 CC properties and impart gloss to the paper during coating step.
 CC The present sequence represents a GTFB mutant of the invention.
 CC Note: The present sequence is not shown in the specification but
 CC was created by the indexer using the GTFB sequence appearing as AA098C27
 CC and the information in claim 36.
 CC Sequence 1475 AA;
 CC SQ

Query Match 92.6%; Score 88; DB 23; Length 1475;
 Best Local Similarity 94.7%; Pred. No. 5.8e-07;
 Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 VPYSFIRADSEVQDLIA 19
 DB 552 VPYSFIRADSEVQDLIA 570
 RESULT 14
 AAU98036
 ID AAU98036 standard; Protein; 1475 AA.
 XX
 AC AAU98036;
 XX
 DT 27-AUG-2002 (first entry)
 XX
 XX S. mutans glucosyltransferase GTFB mutant D567T/D571K.
 DE
 XX Glucosyltransferase; GTFB; transgenic plant; paper sizing;
 KM coating composition; glucan; starch; latex; thermoplastic molecule;
 KM amyloplast; vacuole; paper manufacture; mutant; muten.
 XX
 OS Streptococcus mutans.
 OS Synthetic.
 XX
 XX Key Location/Qualifiers
 FT Misc-difference 567 /note= "Wild-type Asp substituted by Thr"
 FT FT Misc-difference 571 /note= "Wild-type Asp substituted by Lys"
 FT US2002031826-A1.
 XX
 XX 14-MAR-2002.
 XX 19-DEC-2000; 2000US-0740274.
 XX 11-DEC-1998; 98US-0210361.
 XX 07-JUN-1995; 95US-0478704.
 XX 07-JUN-1995; 95US-0482711.
 XX 07-JUN-1995; 95US-0485243.
 XX 16-JAN-1998; 98US-0007999.
 XX 16-JAN-1998; 98US-0008172.
 XX 20-JAN-1998; 98US-0009620.
 XX (NICH/) NICHOLS S E.
 XX Nichols SE;
 XX WPI; 2002-414332/44.
 DR Glucosyltransferase B or D protein useful for producing a glucan useful
 PT as substitutes for and additions to modified starch and latexes in
 PT paper manufacture, comprises mutations in specific positions -
 XX Claim 36; Page -, 44pp; English.
 XX The invention an isolated protein comprising a glucosyltransferase
 CC (GTF) B polypeptide having changes at position from 1448V, D457N,
 CC D567T, K1014T, D457N/D567T, D457N/D571K, D567T/D571K,
 CC D567T/D571K/K1014T, I448V/D457N/D571K/K719Q/K1014T,
 CC Y169A/Y170A/Y171A, and K779Q or a GTF D polypeptide having
 CC changes at positions from T589D, T589E, N471D, N471D/T589D, and
 CC N471D/T589E. Also included are a glucan produced by the GTF mutant,
 CC an isolated polynucleotide which encodes P1 or P2, or its complementary
 CC polynucleotide, a ribonucleic acid sequence encoding the GTF mutant,
 CC an expression cassette comprising the polynucleotide operably linked to a
 CC promoter, a vector comprising the expression cassette, host cell
 CC introduced with the vector, a transgenic plant, a paper sizing and/or
 CC vector, a seed or tuber from the transgenic plant, a paper transformed
 CC coating composition comprising a glucan produced in a plant transformed
 CC with a gene encoding the mutant GTF, wild type or, starch, a latex,
 CC

thermoplastic molecule or their combinations or glucan and starch where the glucan is produced in the amyloplast and/or vacuole or a maize line deficient in starch biosynthesis, transformed with a gene encoding a glucosyltransferase B or D enzyme, wild-type or mutant and a paper comprising the glucan (paper sizing/coating agent). The vector is useful for producing the glucan in a plant. The method comprises transforming a plant cell with the vector, growing the plant cell under plant growing conditions to produce a regenerated plant and inducing expression of the polynucleotide for a time sufficient to produce the glucan in the regenerated plant, where the vector contains a transit sequence from ribulose biphosphate carboxylase small subunit, waxy, brittle-1 and chlorophyll AB binding protein to produce a transgenic plant, and glucan is produced in the amyloplast of potato or the vacuole of sugar beet. Glucans are useful as substitutes for and additions to modified starch and latexes in paper manufacture. Unlike prior art techniques, which require input materials that produce chemical effluents, paper manufacture utilizing the glucan produced by GTF, which utilizes biologically produced input materials, is more cost-effective and environmentally friendly. Moreover, glucans also exhibit thermoplastic properties and impart gloss to the paper during coating step. The present sequence represents a GTFB mutant of the invention. Note: The present sequence is not shown in the specification but was created by the indexer using the GTFB sequence appearing as AAU98027 and the information in claim 36.

Sequence 1475 AA:

Query Match 92.6%; Score 88; DB 23; Length 1475;
Best Local Similarity 94.7%; Pred. No. 5,8e-07;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

1 VPSYSTRAHDSVQDLIA 19
|||||
552 VPSYSTRAHDSVQTLIA 570

Result 15

AAU98037 standard; Protein; 1475 AA.

AAU98037;

27-AUG-2002 (first entry)

S. mutans glucosyltransferase GTFB mutant D567T/D571K/K1014T.

Glucosyltransferase; GTFB; transgenic plant; paper sizing;

coating composition; glucan; starch; latex; thermoplastic molecule;

amyloplast; vacuole; paper manufacture; mutant; mutain.

Streptococcus mutans.

Synthetic.

Key Location/Qualifiers

Misc-difference 567 /note= "Wild-type Asp substituted by Thr"

Misc-difference 571 /note= "Wild-type Asp substituted by Lys"

Misc-difference 1014 /note= "Wild-type Lys substituted by Thr"

US2002031826-A1.

14-MAR-2002.

19-DEC-2000; 2000US-0704274.

11-DEC-1998; 98US-0210361.

07-JUN-1995; 95US-0482704.

07-JUN-1995; 95US-0482711.

07-JUN-1995; 95US-0485243.

16-JAN-1998; 98US-0007999.

16-JAN-1998; 98US-0008172.

PR 20-JAN-1998; 98US-0009620.

XX (NICH/) NICHOLS S E.

XX Nichols SE;

XX WPI; 2002-414332/44.

PT Glucosyltransferase B or D protein useful for producing a glucan useful as substitutes for and additions to modified starch and latexes in

PT paper manufacture, comprises mutations in specific positions

PS Claim 36; Page -; 44pp; English.

The invention an isolated protein comprising a glucosyltransferase (GTF) B polypeptide having changes at position from 1448V, D457N, D567T, K1014T, D457N/D571K, D457N/D571K, D567T/D571K, D567T/D571K/K1014T, 1448V/D457N/D571K/K779Q/K1014T, Y169A/Y170A/Y171A, and K779Q or a GTF D polypeptide having changes at positions from T589D, T589E, N471D, N471D/T589D, and N471D/T589E. Also included are a glucan produced by the GTF mutant, an isolated polynucleotide which encodes P1 or P2, or its complementary polynucleotide, a ribonucleic acid sequence encoding the GTF mutant, an expression cassette comprising the polynucleotide operably linked to a promoter, a vector comprising the expression cassette, host cell introduced with the vector, a transgenic plant comprising the vector, a seed or tuber from the transgenic plant, a paper sizing and/or coating composition comprising a glucan produced in a plant transformed with a gene encoding the mutant GTF, wild type or, starch, a latex, thermoplastic molecule or their combinations or glucan and starch where the glucan is produced in the amyloplast and/or vacuole or a maize line deficient in starch biosynthesis, transformed with a gene encoding a glucosyltransferase B or D enzyme, wild-type or mutant and a paper comprising the glucan (paper sizing/coating agent). The vector is useful for producing a glucan in a plant. The method comprises transforming a plant cell with the vector, growing the plant cell under plant growing conditions to produce a regenerated plant and inducing expression of the polynucleotide for a time sufficient to produce the glucan in the regenerated plant, where the vector contains a transit sequence from ribulose biphosphate carboxylase small subunit, waxy, brittle-1 and chlorophyll AB binding protein to produce a transgenic plant, and glucan is produced in the amyloplast of potato or the vacuole of sugar beet. Glucans are useful as substitutes for and additions to modified starch and latexes in paper manufacture. Unlike prior art techniques, which require input materials that produce chemical effluents, paper manufacture utilizing the glucan produced by GTF, which utilizes biologically produced input materials, is more cost-effective and environmentally friendly. Moreover, glucans also exhibit thermoplastic properties and impart gloss to the paper during coating step. The present sequence represents a GTFB mutant of the invention. Note: The present sequence is not shown in the specification but was created by the indexer using the GTFB sequence appearing as AAU98027 and the information in claim 36.

Sequence 1475 AA:

Query Match 92.6%; Score 88; DB 23; Length 1475;
Best Local Similarity 94.7%; Pred. No. 5,8e-07;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

1 VPSYSTRAHDSVQDLIA 19
|||||
552 VPSYSTRAHDSVQTLIA 570

Search completed: November 13, 2003, 09:38:25
Job time : 34.8463 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 13, 2003, 09:31:40 : Search time 12.365 Seconds
(Without alignments)
65.165 Million cell updates/sec

Title: US-09-290-049A-2

Sequence: 1 VPSYSFIRAHDSVQDLIA 19

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:*

- 1: /cgn2_6/ptodata/2/1aa/5A_COMB.pep:*
- 2: /cgn2_6/ptodata/2/1aa/5B_COMB.pep:*
- 3: /cgn2_6/ptodata/2/1aa/6A_COMB.pep:*
- 4: /cgn2_6/ptodata/2/1aa/6B_COMB.pep:*
- 5: /cgn2_6/ptodata/2/1aa/PTCUTS_COMB.pep:*
- 6: /cgn2_6/ptodata/2/1aa/backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	95	100.0	1475	3	US-09-007-999-2
2	95	100.0	1475	3	US-09-210-361-2
3	95	100.0	1475	4	US-09-740-274-2
4	91	95.8	1375	3	US-09-210-361-4
5	91	95.8	1375	4	US-09-740-274-4
6	80	84.2	523	4	US-09-604-957-5
7	67	70.5	584	4	US-09-604-957-6
8	66	69.5	535	4	US-09-604-957-7
9	66	69.5	1278	4	US-09-604-957-3
10	66	69.5	2057	4	US-09-499-203-2
11	65	68.4	545	4	US-09-604-957-4
12	65	68.4	1430	3	US-09-008-117-2
13	65	68.4	1430	3	US-09-210-361-6
14	65	68.4	1430	4	US-09-740-274-6
15	61	64.2	1577	2	US-08-793-824-2
16	43	44.3	75	4	US-09-107-532A-7020
17	42	44.2	565	4	US-09-107-532A-4217
18	42	44.2	616	4	US-09-268-347-38
19	42	44.2	659	4	US-09-268-347-46
20	42	44.2	660	4	US-09-268-347-45
21	42	44.2	2627	2	US-08-751-189-3
22	42	44.2	2627	2	US-09-060-836-3
23	42	44.2	2627	3	US-09-184-445-3
24	41.5	43.7	713	1	US-08-849-212-4
25	40	42.1	431	1	US-07-783-705A-2
26	39	41.1	309	4	US-09-345-473E-37
27	38.5	40.5	237	2	US-08-578-709-11

28	38.5	40.5	345	3	US-09-231-529-1	Sequence 1, Appli
29	38.5	40.5	345	3	US-08-977-816-1	Sequence 1, Appli
30	38.5	40.5	350	4	US-09-533-494A-17	Sequence 17, Appli
31	38.5	40.5	417	4	US-09-533-494A-13	Sequence 13, Appli
32	38.5	40.5	500	2	US-08-578-709-15	Sequence 15, Appli
33	38	40.0	184	1	US-08-353-550-2	Sequence 2, Appli
34	38	40.0	184	1	US-08-551-687-2	Sequence 2, Appli
35	38	40.0	280	4	US-09-252-991A-18490	Sequence 18490, A
36	38	40.0	331	4	US-09-252-991A-18759	Sequence 18759, A
37	38	40.0	333	4	US-09-252-991A-33113	Sequence 33113, A
38	38	40.0	325	4	US-09-107-532A-6786	Sequence 6786, Ap
39	38	40.0	387	4	US-09-134-001C-3206	Sequence 3206, Ap
40	38	40.0	389	4	US-09-107-532A-6456	Sequence 6456, Ap
41	38	40.0	428	1	US-08-353-550-1	Sequence 1, Appli
42	38	40.0	428	2	US-08-551-687-1	Sequence 1, Appli
43	38	40.0	431	3	US-08-807-342B-5	Sequence 5, Appli
44	38	40.0	469	1	US-08-353-550-6	Sequence 6, Appli
45	38	40.0	469	2	US-08-551-687-6	Sequence 6, Appli

ALIGNMENTS

```

RESULT 1
US-09-007-999-2
Sequence 2, Application US/09007999
Patent No. 6087559
GENERAL INFORMATION:
APPLICANT: Nicholas, Scott E.
TITLE OF INVENTION: Substitutes for Modified Starch and
FILE REFERENCE: 0356D
CURRENT APPLICATION NUMBER: US/09/007,999
EARLIER FILING DATE: 1998-01-16
EARLIER APPLICATION NUMBER: 08/478,704
EARLIER FILING DATE: 1995-06-07
NUMBER OF SEQ ID NOS: 2
SOFTWARE: FASTSEQ for Windows Version 3.0
SEQ ID NO 2
LENGTH: 1475
TYPE: PRT
ORGANISM: Streptococcus mutans
US-09-007-999-2

Query Match      100.0%; Score 95; DB 3; Length 1475;
Best Local Similarity 100.0%; Pred. No. 1.5e-08;
Matches 19; Conservative 0; Mismatches 0; Indels 0;

Cy      1 VPSYSFIRAHDSVQDLIA 19
Db      552 VPSYSFIRAHDSVQDLIA 570

RESULT 2
US-09-210-361-2
Sequence 2, Application US/09210361
Patent No. 6284479
GENERAL INFORMATION:
APPLICANT: Nicholas, Scott E.
TITLE OF INVENTION: Substitutes for Modified Starches and
FILE REFERENCE: 0357CR
CURRENT APPLICATION NUMBER: US/09/210,361
EARLIER FILING DATE: 1998-12-11
EARLIER APPLICATION NUMBER: 09/007,999
EARLIER FILING DATE: 1998-01-16
EARLIER APPLICATION NUMBER: 08/478,704
EARLIER FILING DATE: 1995-06-07
EARLIER APPLICATION NUMBER: 09/009,620
EARLIER FILING DATE: 1998-01-20
EARLIER APPLICATION NUMBER: 08/485,243
EARLIER FILING DATE: 1995-06-07
EARLIER APPLICATION NUMBER: 09/008,172

```

```

; EARLIER FILING DATE: 1998-01-16
; EARLIER APPLICATION NUMBER: 08/482,711
; EARLIER FILING DATE: 1995-06-07
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 1475
; TYPE: PRT
; ORGANISM: Streptococcus mutans
US-09-210-361-2

Query Match          100.0%; Score 95; DB 3; Length 1475;
Best Local Similarity 100.0%; Pred. No. 1.5e-08;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 VPSYSFIRAHDSVQDLIA 19
Db      552 VPSYSFIRAHDSVQDLIA 570

RESULT 3
US-09-740-274-2
; Sequence 2, Application US/09740274
; Patent No. 6465203
; GENERAL INFORMATION:
; APPLICANT: Nicholas, Scott E.
; TITLE OF INVENTION: Glucan-containing Compositions and Paper
; FILE REFERENCE: 0357CRD
; CURRENT APPLICATION NUMBER: US/09/740,274
; PRIOR FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: 09/210,361
; PRIOR FILING DATE: 1998-12-11
; PRIOR APPLICATION NUMBER: 09/007,999
; PRIOR FILING DATE: 1998-01-16
; PRIOR APPLICATION NUMBER: 08/478,704
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: 09/009,620
; PRIOR FILING DATE: 1998-01-20
; PRIOR APPLICATION NUMBER: 08/485,243
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: 09/008,172
; PRIOR FILING DATE: 1998-01-16
; PRIOR APPLICATION NUMBER: 08/482,711
; PRIOR FILING DATE: 1995-06-07
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 1475
; TYPE: PRT
; ORGANISM: Streptococcus mutans
US-09-740-274-2

Query Match          100.0%; Score 95; DB 4; Length 1475;
Best Local Similarity 100.0%; Pred. No. 1.5e-08;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 VPSYSFIRAHDSVQDLIA 19
Db      552 VPSYSFIRAHDSVQDLIA 570

RESULT 4
US-09-210-361-4
; Sequence 4, Application US/09210361
; Patent No. 6284479
; GENERAL INFORMATION:
; APPLICANT: Nicholas, Scott E.
; TITLE OF INVENTION: Substitutes for Modified Starches and
; FILE REFERENCE: 0357CR
; CURRENT APPLICATION NUMBER: US/09/210,361
; PRIOR FILING DATE: 1998-12-11
; EARLIER APPLICATION NUMBER: 09/007,999
```

```

; EARLIER FILING DATE: 1998-01-16
; EARLIER APPLICATION NUMBER: 08/478,704
; EARLIER FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 09/009,620
; EARLIER FILING DATE: 1998-01-20
; EARLIER APPLICATION NUMBER: 08/485,243
; EARLIER FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 09/008,172
; EARLIER FILING DATE: 1998-01-16
; EARLIER APPLICATION NUMBER: 08/482,711
; EARLIER FILING DATE: 1995-06-07
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 1375
; TYPE: PRT
; ORGANISM: streptococcus mutans
US-09-740-274-4

Query Match          95.8%; Score 91; DB 3; Length 1375;
Best Local Similarity 100.0%; Pred. No. 7.3e-08;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 VPSYSFIRAHDSVQDLI 18
Db      578 VPSYSFIRAHDSVQDLI 595

RESULT 5
US-09-740-274-4
; Sequence 4, Application US/09740274
; Patent No. 6465203
; GENERAL INFORMATION:
; APPLICANT: Nicholas, Scott E.
; TITLE OF INVENTION: Glucan-containing Compositions and Paper
; FILE REFERENCE: 0357CRD
; CURRENT APPLICATION NUMBER: US/09/740,274
; PRIOR FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: 09/210,361
; PRIOR FILING DATE: 1998-12-11
; PRIOR APPLICATION NUMBER: 09/007,999
; PRIOR FILING DATE: 1998-01-16
; PRIOR APPLICATION NUMBER: 08/478,704
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: 09/009,620
; PRIOR FILING DATE: 1998-01-20
; PRIOR APPLICATION NUMBER: 08/485,243
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: 09/008,172
; PRIOR FILING DATE: 1998-01-16
; PRIOR APPLICATION NUMBER: 08/482,711
; PRIOR FILING DATE: 1995-06-07
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 1375
; TYPE: PRT
; ORGANISM: streptococcus mutans
US-09-740-274-4

Query Match          95.8%; Score 91; DB 4; Length 1375;
Best Local Similarity 100.0%; Pred. No. 7.3e-08;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 VPSYSFIRAHDSVQDLI 18
Db      578 VPSYSFIRAHDSVQDLI 595

RESULT 6
US-09-604-957-5
; Sequence 5, Application US/09604957
; Patent No. 6486314
```

```
; GENERAL INFORMATION:
; APPLICANT: VAN GEEL-SCHUTTEN, GERRITDINA HENDRIKA
; APPLICANT: DIJKHUIZEN, LUBBERT
; APPLICANT: RAHAOUI, HAKIM
; APPLICANT: LEER, ROBERT-JAN
; TITLE OF INVENTION: NOVEL GLUCOSYLTRANSFERASE AND NOVEL GLUCAN
; FILE REFERENCE: BO 43388
; CURRENT APPLICATION NUMBER: US/09/604,957
; CURRENT FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: 00201871.1
; PRIOR FILING DATE: 2000-05-25
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 523
; TYPE: PRT
; ORGANISM: Leuconostoc mesenteroides
; US-09-604-957-5

Query Match          84.2%; Score 80; DB 4; Length 523;
Best Local Similarity 73.7%; Pred. No. 2,3e-06;
Matches 14; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY      1 VPSYSFIRAHDSVODLIA 19
Db      146 IPNYSFVRADHSDVOTVIA 164

RESULT 7
US-09-604-957-6
; Sequence 6, Application US/09604957
; Patent No. 6486314
; GENERAL INFORMATION:
; APPLICANT: VAN GEEL-SCHUTTEN, GERRITDINA HENDRIKA
; APPLICANT: DIJKHUIZEN, LUBBERT
; APPLICANT: RAHAOUI, HAKIM
; APPLICANT: LEER, ROBERT-JAN
; TITLE OF INVENTION: NOVEL GLUCOSYLTRANSFERASE AND NOVEL GLUCAN
; FILE REFERENCE: BO 43388
; CURRENT APPLICATION NUMBER: US/09/604,957
; CURRENT FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: 00201871.1
; PRIOR FILING DATE: 2000-05-25
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 584
; TYPE: PRT
; ORGANISM: Leuconostoc mesenteroides
; US-09-604-957-6

Query Match          70.5%; Score 67; DB 4; Length 584;
Best Local Similarity 57.9%; Pred. No. 0.00062;
Matches 11; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

QY      1 VPSYSFIRAHDSVODLIA 19
Db      167 IPNYSFVRADHDADDPIS 185

RESULT 8
US-09-604-957-7
; Sequence 7, Application US/09604957
; Patent No. 6486314
; GENERAL INFORMATION:
; APPLICANT: VAN GEEL-SCHUTTEN, GERRITDINA HENDRIKA
; APPLICANT: DIJKHUIZEN, LUBBERT
; APPLICANT: RAHAOUI, HAKIM
; APPLICANT: LEER, ROBERT-JAN
; TITLE OF INVENTION: NOVEL GLUCOSYLTRANSFERASE AND NOVEL GLUCAN
; FILE REFERENCE: BO 43388
; CURRENT APPLICATION NUMBER: US/09/604,957
; CURRENT FILING DATE: 2000-06-28
```

```
; PRIOR APPLICATION NUMBER: 00201871.1
; PRIOR FILING DATE: 2000-05-25
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 7
; LENGTH: 535
; TYPE: PRT
; ORGANISM: Lactobacillus reuteri
; US-09-604-957-7

Query Match          69.5%; Score 66; DB 4; Length 535;
Best Local Similarity 61.1%; Pred. No. 0.00084;
Matches 11; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY      1 VPSYSFIRAHDSVODLI 18
Db      144 IPNYSFVRADHNSQDQI 161

RESULT 9
US-09-604-957-3
; Sequence 3, Application US/09604957
; Patent No. 6486314
; GENERAL INFORMATION:
; APPLICANT: VAN GEEL-SCHUTTEN, GERRITDINA HENDRIKA
; APPLICANT: DIJKHUIZEN, LUBBERT
; APPLICANT: RAHAOUI, HAKIM
; APPLICANT: LEER, ROBERT-JAN
; TITLE OF INVENTION: NOVEL GLUCOSYLTRANSFERASE AND NOVEL GLUCAN
; FILE REFERENCE: BO 43388
; CURRENT APPLICATION NUMBER: US/09/604,957
; CURRENT FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: 00201871.1
; PRIOR FILING DATE: 2000-05-25
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 1278
; TYPE: PRT
; ORGANISM: Lactobacillus reuteri
; US-09-604-957-3

Query Match          69.5%; Score 66; DB 4; Length 1278;
Best Local Similarity 61.1%; Pred. No. 0.0024;
Matches 11; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY      1 VPSYSFIRAHDSVODLI 18
Db      620 IPNYSFVRADHNSQDQI 637

RESULT 10
US-09-499-203-2
; Sequence 2, Application US/09499203
; Patent No. 6570065
; GENERAL INFORMATION:
; APPLICANT: KOSSMANN, Jens
; APPLICANT: WELSH, Thomas
; APPLICANT: CUANZ, Martin
; APPLICANT: KNUTH, Karola
; TITLE OF INVENTION: Nucleic Acid Molecules Encoding Alternansucrase
; FILE REFERENCE: 147-196P
; CURRENT APPLICATION NUMBER: US/09/499,203
; CURRENT FILING DATE: 2000-02-08
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 2057
; TYPE: PRT
; ORGANISM: Leuconostoc mesenteroides
; US-09-499-203-2

Query Match          69.5%; Score 66; DB 4; Length 2057;
```

Best Local Similarity 61.1%; Pred. No. 0.0043;
Matches 11; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 VPSFIRAHSEVODLIA 18
:|||||:|
Db 757 INYSFVRHVDVADQPEI 774

RESULT 11
US-09-604-957-4
; Sequence 4, Application US/09604957
; Patent No. 6486314
; GENERAL INFORMATION:
; APPLICANT: VAN GEEL-SCHUTTEN, GERRITDINA HENDRIKA
; APPLICANT: DIJKHUIZEN, LOBBERT
; APPLICANT: RAHMOUI, HAKIM
; APPLICANT: LEER, ROBERT-JAN
; TITLE OF INVENTION: NOVEL GLUCOSYLTRANSFERASE AND NOVEL GLUCAN
; FILE REFERENCE: BO 43388
; CURRENT APPLICATION NUMBER: US/09/604,957
; CURRENT FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: 00201871.1
; PRIOR FILING DATE: 2000-05-25
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 545
; TYPE: PRT
; ORGANISM: Streptococcus mutans
US-09-604-957-4

Query Match 68.4%; Score 65; DB 4; Length 545;
Best Local Similarity 76.5%; Pred. No. 0.0013; 2; Indels 0; Gaps 0;
Matches 13; Conservative 2; Mismatches 2;

QY 3 SYSFIRAHSEVODLIA 19
:|||||:|
Db 156 NYFIRAHSEVQTVIA 172

RESULT 12
US-09-008-172-2
; Sequence 2, Application US/09008172
; Patent No. 6127602
; GENERAL INFORMATION:
; APPLICANT: NICHOLS, Scott E.
; TITLE OF INVENTION: Substitutes for Modified Starch and
; FILE REFERENCE: 0358D
; CURRENT APPLICATION NUMBER: US/09/008,172
; CURRENT FILING DATE: 1998-01-16
; EARLIER APPLICATION NUMBER: 08/482,711
; EARLIER FILING DATE: 1995-06-07
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 1430
; TYPE: PRT
; ORGANISM: Streptococcus mutans
US-09-008-172-2

Query Match 68.4%; Score 65; DB 3; Length 1430;
Best Local Similarity 76.5%; Pred. No. 0.0042; 2; Indels 0; Gaps 0;
Matches 13; Conservative 2; Mismatches 2;

QY 3 SYSFIRAHSEVODLIA 19
:|||||:|
Db 576 NYFIRAHSEVQTVIA 592

RESULT 13
US-09-210-361-6
; Sequence 6, Application US/09210361

; Patent No. 6284479
; GENERAL INFORMATION:
; APPLICANT: NICHOLS, Scott E.
; TITLE OF INVENTION: Substitutes for Modified Starches and
; FILE REFERENCE: 0357CR
; CURRENT APPLICATION NUMBER: US/09/210,361
; CURRENT FILING DATE: 1998-12-11
; EARLIER APPLICATION NUMBER: 09/007,999
; EARLIER FILING DATE: 1998-01-16
; EARLIER APPLICATION NUMBER: 08/478,704
; EARLIER FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 09/009,620
; EARLIER FILING DATE: 1998-01-20
; EARLIER APPLICATION NUMBER: 08/485,243
; EARLIER FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 09/008,172
; EARLIER FILING DATE: 1998-01-16
; EARLIER APPLICATION NUMBER: 08/482,711
; EARLIER FILING DATE: 1995-06-07
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 6
; LENGTH: 1430
; TYPE: PRT
; ORGANISM: streptococcus mutans
US-09-210-361-6

Query Match 68.4%; Score 65; DB 3; Length 1430;
Best Local Similarity 76.5%; Pred. No. 0.0042; 2; Indels 0; Gaps 0;
Matches 13; Conservative 2; Mismatches 2;

QY 3 SYSFIRAHSEVODLIA 19
:|||||:|
Db 576 NYFIRAHSEVQTVIA 592

RESULT 14
US-09-740-274-6
; Sequence 6, Application US/09740274
; Patent No. 6465203
; GENERAL INFORMATION:
; APPLICANT: NICHOLS, Scott E.
; TITLE OF INVENTION: Glucan-containing Compositions and Paper
; FILE REFERENCE: 0357CRD
; CURRENT APPLICATION NUMBER: US/09/740,274
; CURRENT FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: 09/210,361
; PRIOR FILING DATE: 1998-12-11
; PRIOR APPLICATION NUMBER: 09/007,999
; PRIOR FILING DATE: 1998-01-16
; PRIOR APPLICATION NUMBER: 08/478,704
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: 09/009,620
; PRIOR FILING DATE: 1998-01-20
; PRIOR APPLICATION NUMBER: 08/485,243
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: 09/008,172
; PRIOR FILING DATE: 1998-01-16
; PRIOR APPLICATION NUMBER: 08/482,711
; PRIOR FILING DATE: 1995-06-07
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 6
; LENGTH: 1430
; TYPE: PRT
; ORGANISM: streptococcus mutans
US-09-740-274-6

Query Match 68.4%; Score 65; DB 4; Length 1430;
Best Local Similarity 76.5%; Pred. No. 0.0042; 2; Indels 0; Gaps 0;
Matches 13; Conservative 2; Mismatches 2;

QY 3 SYSFIRAHDESEVODLIA 19
 :|||||:|
 Db 576 NYIFIRAHDESEVQAVIA 592

RESULT 15

US-08-793-824-2
 ; Sequence 2, Application US/08793824
 ; Patent No. 5981838
 ; GENERAL INFORMATION:
 ; APPLICANT: Simpson, Christine Lynn
 ; APPLICANT: Giffard, Philip Morrison
 ; APPLICANT: Jacques, Nicholas Anthony
 ; TITLE OF INVENTION: Genetic Manipulation of Plants to
 ; TITLE OF INVENTION: Increase Stored Carbohydrates
 ; NUMBER OF SEQUENCES: 2
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Griffith Hack & Co
 ; STREET: Level 8, 168 Walker Street
 ; CITY: No. 5981838th, Sydney
 ; STATE: New South Wales
 ; COUNTRY: Australia
 ; ZIP: 2060
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patentln Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/793,824
 ; FILING DATE:
 ; CLASSIFICATION: 800
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: AU PM7643
 ; FILING DATE: 24-AUG-1994
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 61 2 9957 5944
 ; TELEFAX: 61 2 957 6288
 ; TELEX: 26547
 ; INFORMATION FOR SEQ ID NO: 2:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1577 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS:
 ; TOPOLOGY: not relevant
 ; MOLECULE TYPE: protein
 ; ORIGINAL SOURCE:
 ; ORGANISM: Streptococcus salivarius
 ; US-08-793-824-2

Query Match 64.2%; Score 61; DB 2; Length 1577;
 Best Local Similarity 64.7%; Pred. No. 0.025;
 Matches 11; Conservative 4; Mismatches 2; Indels 0; Gaps 0;
 QY 3 SYSFIRAHDESEVODLIA 19
 :|||||:|
 Db 661 NYIFIRAHDESEVQAVIA 677

Search completed: November 13, 2003, 09:47:57
 Job time : 12.3365 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 13, 2003, 09:45:40 : Search time 24.8531 Seconds
(without alignments)
139.565 Million cell updates/sec

Title: US-09-290-049a-2

Perfect score: 95

Sequence: 1 VPSSYFIRADSEVQDLIA 19

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 666188 seqs, 182559486 residues

Total number of hits satisfying chosen parameters: 666188

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

1: /cgn2_6/ptodata/2/pubppaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/2/pubppaa/US06_PUBCOMB.pep.*
3: /cgn2_6/ptodata/2/pubppaa/US06_PUBCOMB.pep.*
4: /cgn2_6/ptodata/2/pubppaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubppaa/US06_PUBCOMB.pep.*
6: /cgn2_6/ptodata/2/pubppaa/US06_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubppaa/US06_PUBCOMB.pep.*
8: /cgn2_6/ptodata/2/pubppaa/US06_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubppaa/US06_PUBCOMB.pep.*
10: /cgn2_6/ptodata/2/pubppaa/US06_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubppaa/US06_PUBCOMB.pep.*
12: /cgn2_6/ptodata/2/pubppaa/US06_PUBCOMB.pep.*
13: /cgn2_6/ptodata/2/pubppaa/US06_PUBCOMB.pep.*
14: /cgn2_6/ptodata/2/pubppaa/US06_PUBCOMB.pep.*
15: /cgn2_6/ptodata/2/pubppaa/US06_PUBCOMB.pep.*
16: /cgn2_6/ptodata/2/pubppaa/US06_PUBCOMB.pep.*
17: /cgn2_6/ptodata/2/pubppaa/US06_PUBCOMB.pep.*
18: /cgn2_6/ptodata/2/pubppaa/US06_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	95	100.0	1475	9 US-09-740-274-2	Sequence 2, Appl1
2	91	95.8	1375	9 US-09-740-274-4	Sequence 11, Appl1
3	80	84.2	522	10 US-09-995-749A-11	Sequence 13, Appl1
4	66	69.5	535	10 US-09-995-749A-13	Sequence 12, Appl1
5	66	69.5	584	10 US-09-995-749A-12	Sequence 10, Appl1
6	66	69.5	1781	10 US-09-995-749A-2	Sequence 6, Appl1
7	65	68.4	545	10 US-09-995-749A-10	Sequence 10, Appl1
8	65	68.4	1430	9 US-09-740-274-6	Sequence 5098, Ap
9	42	44.2	115	11 US-09-764-891-5098	Sequence 7564, Ap
10	42	44.2	364	15 US-10-156-761-7564	Sequence 31, Appl1
11	41	43.2	170	9 US-09-775-925-31	Sequence 362, App
12	41	43.2	458	15 US-10-103-813-362	Sequence 2, Appl1
13	41	43.2	552	10 US-09-775-925-2	Sequence 15, Appl1
14	41	43.2	552	10 US-09-955-732-15	Sequence 6490, Ap
15	41	43.2	662	10 US-09-738-626-6490	

16	40.5	42.6	296	14 US-10-080-960-19	Sequence 19, Appl1
17	40	42.1	170	9 US-08-775-925-82	Sequence 32, Appl1
18	40	42.1	241	9 US-09-775-925-8	Sequence 8, Appl1
19	40	42.1	484	12 US-10-181-590-9	Sequence 9, Appl1
20	40	42.1	509	9 US-09-775-925-6	Sequence 6, Appl1
21	40	42.1	509	10 US-09-955-732-14	Sequence 14, Appl1
22	40	42.1	547	11 US-09-943-857-2	Sequence 2, Appl1
23	40	42.1	547	11 US-09-943-857-4	Sequence 4, Appl1
24	40	42.1	547	11 US-09-943-857-6	Sequence 6, Appl1
25	40	42.1	572	9 US-09-815-242-5626	Sequence 5626, Ap
26	40	42.1	572	9 US-09-815-242-12247	Sequence 12247, A
27	40	42.1	573	9 US-09-815-242-12921	Sequence 12921, A
28	39	41.1	67	15 US-10-007-280A-214	Sequence 214, App
29	39	41.1	309	10 US-09-862-027-37	Sequence 37, Appl1
30	39	41.1	405	15 US-10-156-761-9148	Sequence 9148, Ap
31	39	41.1	489	10 US-09-738-626-3805	Sequence 3805, Ap
32	39	41.1	557	11 US-09-934-455-326	Sequence 326, App
33	39	41.1	779	12 US-10-353-929-49	Sequence 82, Appl1
34	39	41.1	1383	15 US-10-021-955-82	Sequence 86, Appl1
35	39	41.1	1383	15 US-10-021-955-86	Sequence 87, Appl1
36	39	41.1	1383	15 US-10-021-955-87	Sequence 79, Appl1
37	39	41.1	1389	15 US-10-021-955-79	Sequence 45372, A
38	38	40.0	166	9 US-09-864-761-45372	Sequence 5731, Ap
39	38	40.0	312	10 US-09-738-626-5731	Sequence 12045, A
40	38	40.0	421	15 US-10-156-761-12914	Sequence 3363, Ap
41	38	40.0	574	9 US-09-815-242-12045	Sequence 2, Appl1
42	38	40.0	831	15 US-10-128-714-3363	Sequence 7, Appl1
43	38	40.0	1025	15 US-10-055-475-2	Sequence 9, Appl1
44	38	40.0	1025	15 US-10-055-475-7	
45	38	40.0	1025	15 US-10-055-475-9	

ALIGNMENTS

RESULT 1
US-09-740-274-2
Sequence 2, Application US/09740274
Patent No. US20020031826A1
GENERAL INFORMATION:
APPLICANT: Nichols, Scott E.
TITLE OF INVENTION: Glucan-containing Compositions and Paper
FILE REFERENCE: 0357CRD
CURRENT FILING DATE: 1998-12-11
CURRENT APPLICATION NUMBER: US/09/740,274
PRIOR FILING DATE: 1998-12-11
PRIOR APPLICATION NUMBER: 09/210,361
PRIOR FILING DATE: 1998-01-16
PRIOR APPLICATION NUMBER: 08/478,704
PRIOR FILING DATE: 1995-06-07
PRIOR APPLICATION NUMBER: 09/009,620
PRIOR FILING DATE: 1998-01-20
PRIOR APPLICATION NUMBER: 08/485,243
PRIOR FILING DATE: 1995-06-07
PRIOR APPLICATION NUMBER: 09/008,172
PRIOR FILING DATE: 1998-01-16
PRIOR APPLICATION NUMBER: 08/482,711
PRIOR FILING DATE: 1995-06-07
NUMBER OF SEQ ID NOS: 6
SOFTWARE: FastrSeq for Windows Version 3.0
SEQ ID NO 2
LENGTH: 1475
TYPE: PRT
ORGANISM: Streptococcus mutans
US-09-740-274-2

Query Match 100.0%; Score 95; DB 9; Length 1475;
Best Local Similarity 100.0%; Pred. No. 1.7e-07;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 VPSSYFIRADSEVQDLIA 19
|||||

Db 552 VPSYSTRADHSEVQDLIA 570

RESULT 2
US-09-740-274-4
; Sequence 4, Application US/09740274
; Patent No. US20020031826A1
; GENERAL INFORMATION:
; APPLICANT: Nichols, Scott E.
; TITLE OF INVENTION: Glucan-containing Compositions and Paper
; FILE REFERENCE: 0357CRD
; CURRENT APPLICATION NUMBER: US/09/740,274
; CURRENT FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: 09/210,361
; PRIOR FILING DATE: 1998-12-11
; PRIOR APPLICATION NUMBER: 09/007,999
; PRIOR FILING DATE: 1998-01-16
; PRIOR APPLICATION NUMBER: 08/478,704
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: 09/009,620
; PRIOR FILING DATE: 1998-01-20
; PRIOR APPLICATION NUMBER: 08/485,243
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: 09/008,172
; PRIOR FILING DATE: 1998-01-16
; PRIOR APPLICATION NUMBER: 08/482,711
; PRIOR FILING DATE: 1995-06-07
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 1375
; TYPE: PRT
; ORGANISM: streptococcus mutans
US-09-740-274-4

Query Match 95.8%; Score 91; DB 9; Length 1375;
Best Local Similarity 100.0%; Pred. No. 7.9e-07;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VPSYSTRADHSEVQDLI 18
Db 578 VPSYSTRADHSEVQDLI 595

RESULT 3
US-09-995-749A-11
; Sequence 11, Application US/09995749A
; Patent No. US2002015568A1
; GENERAL INFORMATION:
; APPLICANT: VAN GEEL-SCHUTTEN, GERRITDINA HENDRIKA
; APPLICANT: DIJKHUIZEN, LUBBERT
; APPLICANT: RAHMOUI, HAKIM
; APPLICANT: LEER, ROBERT-JAN
; TITLE OF INVENTION: NOVEL GLUCOSYLTRANSFERASES
; FILE REFERENCE: B043388-CIP
; CURRENT APPLICATION NUMBER: US/09/995,749A
; CURRENT FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: 09/604,957
; PRIOR FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: EPO 00201871.1
; PRIOR FILING DATE: 2000-05-25
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 11
; LENGTH: 522
; TYPE: PRT
; ORGANISM: leuconostoc mesenteroides
US-09-995-749A-11

Query Match 84.2%; Score 80; DB 10; Length 522;
Best Local Similarity 73.7%; Pred. No. 2.1e-05;
Matches 14; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 VPSYSTRADHSEVQDLIA 19
Db 146 IPNYSFVRADHSEVQTVIA 164

RESULT 4
US-09-995-749A-13
; Sequence 13, Application US/09995749A
; Patent No. US2002015568A1
; GENERAL INFORMATION:
; APPLICANT: VAN GEEL-SCHUTTEN, GERRITDINA HENDRIKA
; APPLICANT: DIJKHUIZEN, LUBBERT
; APPLICANT: RAHMOUI, HAKIM
; APPLICANT: LEER, ROBERT-JAN
; TITLE OF INVENTION: NOVEL GLUCOSYLTRANSFERASES
; FILE REFERENCE: B043388-CIP
; CURRENT APPLICATION NUMBER: US/09/995,749A
; CURRENT FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: 09/604,957
; PRIOR FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: EPO 00201871.1
; PRIOR FILING DATE: 2000-05-25
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 13
; LENGTH: 535
; TYPE: PRT
; ORGANISM: lactobacillus reuteri
US-09-995-749A-13

Query Match 69.5%; Score 66; DB 10; Length 535;
Best Local Similarity 61.1%; Pred. No. 0.0057;
Matches 11; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 VPSYSTRADHSEVQDLI 18
Db 144 IPNYSFVRADHSDNDSQDPI 161

RESULT 5
US-09-995-749A-12
; Sequence 12, Application US/09995749A
; Patent No. US2002015568A1
; GENERAL INFORMATION:
; APPLICANT: VAN GEEL-SCHUTTEN, GERRITDINA HENDRIKA
; APPLICANT: DIJKHUIZEN, LUBBERT
; APPLICANT: RAHMOUI, HAKIM
; APPLICANT: LEER, ROBERT-JAN
; TITLE OF INVENTION: NOVEL GLUCOSYLTRANSFERASES
; FILE REFERENCE: B043388-CIP
; CURRENT APPLICATION NUMBER: US/09/995,749A
; CURRENT FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: 09/604,957
; PRIOR FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: EPO 00201871.1
; PRIOR FILING DATE: 2000-05-25
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 12
; LENGTH: 584
; TYPE: PRT
; ORGANISM: leuconostoc mesenteroides
US-09-995-749A-12

Query Match 69.5%; Score 66; DB 10; Length 584;
Best Local Similarity 61.1%; Pred. No. 0.0062;
Matches 11; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 VPSYSTRADHSEVQDLI 18
Db 167 IPNYSFVRADHSDNDSQDPI 184

```

RESULT 6
US-09-995-749A-2
; Sequence 2, Application US/09995749A
; Patent No. US20020155568A1
; GENERAL INFORMATION:
; APPLICANT: VAN GEEL-SCHUTTEN, GERRITDINA HENDRIKA
; APPLICANT: DIJKHUIZEN, LOBBERT
; APPLICANT: RAHAOUI, HAKIM
; APPLICANT: LEER, ROBERT-JAN
; TITLE OF INVENTION: NOVEL GLUCOSYLTRANSFERASES
; FILE REFERENCE: BO43388-CIP
; CURRENT APPLICATION NUMBER: US/09/995,749A
; PRIOR FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: 09/604,957
; PRIOR FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: EPO 00201871.1
; PRIOR FILING DATE: 2000-05-25
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 1781
; TYPE: PRT
; ORGANISM: Lactobacillus reuteri
US-09-995-749A-2

Query Match          69.5%; Score 66; DB 10; Length 1781;
Best Local Similarity 61.1%; Pred. No. 0.022;
Matches 11; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY      1 VPSYGFRAHDSVQDLI 18
DB      1123 INYSFVRADHNSQDQI 1140

RESULT 7
US-09-995-749A-10
; Sequence 10, Application US/09995749A
; Patent No. US20020155568A1
; GENERAL INFORMATION:
; APPLICANT: VAN GEEL-SCHUTTEN, GERRITDINA HENDRIKA
; APPLICANT: DIJKHUIZEN, LOBBERT
; APPLICANT: RAHAOUI, HAKIM
; APPLICANT: LEER, ROBERT-JAN
; TITLE OF INVENTION: NOVEL GLUCOSYLTRANSFERASES
; FILE REFERENCE: BO43388-CIP
; CURRENT APPLICATION NUMBER: US/09/995,749A
; CURRENT FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: 09/604,957
; PRIOR FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: EPO 00201871.1
; PRIOR FILING DATE: 2000-05-25
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
; LENGTH: 545
; TYPE: PRT
; ORGANISM: Streptococcus mutans
US-09-995-749A-10

Query Match          68.4%; Score 65; DB 10; Length 545;
Best Local Similarity 76.5%; Pred. No. 0.0086;
Matches 13; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY      3 SYSFIRAHDSVQDLIA 19
DB      156 NYIFIRAHDSVQTVIA 172

RESULT 8
US-09-740-274-6
; Sequence 6, Application US/09740274
; Patent No. US20020031826A1
; GENERAL INFORMATION:

```

```

; APPLICANT: Nicholas, Scott E.
; TITLE OF INVENTION: Glucan-containing Compositions and Paper
; FILE REFERENCE: 0357CRD
; CURRENT APPLICATION NUMBER: US/09/740,274
; CURRENT FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: 09/210,361
; PRIOR FILING DATE: 1998-12-11
; PRIOR APPLICATION NUMBER: 09/007,999
; PRIOR FILING DATE: 1998-01-16
; PRIOR APPLICATION NUMBER: 08/478,704
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: 09/009,620
; PRIOR FILING DATE: 1998-01-20
; PRIOR APPLICATION NUMBER: 08/485,243
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: 09/008,172
; PRIOR FILING DATE: 1998-01-16
; PRIOR APPLICATION NUMBER: 08/482,711
; PRIOR FILING DATE: 1995-06-07
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 6
; LENGTH: 1430
; TYPE: PRT
; ORGANISM: streptococcus mutans
US-09-740-274-6

Query Match          68.4%; Score 65; DB 9; Length 1430;
Best Local Similarity 76.5%; Pred. No. 0.026;
Matches 13; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY      3 SYSFIRAHDSVQDLIA 19
DB      576 NYIFIRAHDSVQTVIA 592

RESULT 9
US-09-764-891-5098
; Sequence 5098, Application US/09764891
; Publication No. US20030077808A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC006
; CURRENT APPLICATION NUMBER: US/09/764,891
; CURRENT FILING DATE: 2001-01-17
; PRIOR application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 10231
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5098
; LENGTH: 115
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (3)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (9)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (109)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (112)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-764-891-5098

Query Match          44.2%; Score 42; DB 11; Length 115;
Best Local Similarity 38.9%; Pred. No. 14;
Matches 7; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

QY      2 PYSYFRAHDSVQDLIA 19

```

Db 63 PAAVFRHRAQRCQJVA 80

RESULT 10
US-10-156-761-7564
; Sequence 7564, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIRAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 7564
; LENGTH: 364
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-7564

Query Match 44.2%; Score 42; DB 15; Length 364;
Best Local Similarity 44.4%; Pred. No. 51;
Matches 8; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY 2 PSYSFIRAHDESEVODLIA 19
Db 15 PEIKGLRAHDEBAIELIA 32

RESULT 11
US-09-775-925-31
; Sequence 31, Application US/09775925
; Patent No. US20010049358A1
; GENERAL INFORMATION:
; APPLICANT: Lucbe, Ralf M.
; APPLICANT: Wei, Bo
; TITLE OF INVENTION: DSP-12 AND DSP-13 DUAL-SPECIFICITY
; FILE OF INVENTION: PHOSPHATASES
; FILE REFERENCE: 200125.420
; CURRENT APPLICATION NUMBER: US/09/775,925
; CURRENT FILING DATE: 2001-02-01
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 31
; LENGTH: 170
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-775-925-31

Query Match 43.2%; Score 41; DB 9; Length 170;
Best Local Similarity 47.1%; Pred. No. 32;
Matches 8; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 3 SYSFIRAHDESEVODLIA 19
Db 72 AYHNIRVYDEETDILIA 88

RESULT 12
US-10-103-313-362
; Sequence 362, Application US/10103313
; Publication No. US20030082758A1

; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: P1207C1
; CURRENT APPLICATION NUMBER: US/10/103,313
; CURRENT FILING DATE: 2002-03-12
; NUMBER OF SEQ ID NOS: 653
; Prior Application removed - See File Wrapper or Palm
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 362
; LENGTH: 458
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (304)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: misc.feature
; LOCATION: (308)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: misc.feature
; LOCATION: (377)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-10-103-313-362

Query Match 43.2%; Score 41; DB 15; Length 458;
Best Local Similarity 47.1%; Pred. No. 99;
Matches 8; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 3 SYSFIRAHDESEVODLIA 19
Db 41 AYHNIRVYDEETDILIA 57

RESULT 13
US-09-775-925-2
; Sequence 2, Application US/09775925
; Patent No. US20010049358A1
; GENERAL INFORMATION:
; APPLICANT: Lucbe, Ralf M.
; APPLICANT: Wei, Bo
; TITLE OF INVENTION: DSP-12 AND DSP-13 DUAL-SPECIFICITY
; FILE OF INVENTION: PHOSPHATASES
; FILE REFERENCE: 200125.420
; CURRENT APPLICATION NUMBER: US/09/775,925
; CURRENT FILING DATE: 2001-02-01
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 552
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-775-925-2

Query Match 43.2%; Score 41; DB 9; Length 552;
Best Local Similarity 47.1%; Pred. No. 12e+02;
Matches 8; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 3 SYSFIRAHDESEVODLIA 19
Db 214 AYHNIRVYDEETDILIA 230

RESULT 14
US-09-955-732-15
; Sequence 15, Application US/09955732
; Publication No. US20020182203A1
; GENERAL INFORMATION:
; APPLICANT: Lucbe, Ralf M.
; APPLICANT: Wei, Bo
; TITLE OF INVENTION: DSP-15 DUAL-SPECIFICITY PHOSPHATASE
; FILE REFERENCE: 200125.433
; CURRENT APPLICATION NUMBER: US/09/955,732

; CURRENT FILING DATE: 2001-09-18
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15
; LENGTH: 552
; TYPE: PRF
; ORGANISM: Homo sapiens
US-09-955-732-15

Query Match 43.2%; Score 41; DB 10; Length 552;
Best Local Similarity 47.1%; Pred. No. 1.2e+02;
Matches 8; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 3 SYSFIRAHSEVODLIA 19
DB 214 AYHNIRVDEETDILA 230

RESULT 15
US-09-738-626-6490
; Sequence 6490, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAOKO
; APPLICANT: SENOH, AKIHIRO
; APPLICANT: IKEDA, MASAKO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738,626
; CURRENT FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 6490
; LENGTH: 662
; TYPE: PRF
; ORGANISM: Corynebacterium glutamicum
US-09-738-626-6490

Query Match 43.2%; Score 41; DB 10; Length 662;
Best Local Similarity 53.3%; Pred. No. 1.5e+02;
Matches 8; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 5 SFIRAHSEVODLIA 19
DB 415 SIEKLHDSPAODLVA 429

Search completed: November 13, 2003, 10:29:00
Job time : 25.8531 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 13, 2003, 09:31:40 : Search time 6.8436 Seconds
(without alignments)
130.561 Million cell updates/sec

Title: US-09-290-049a-2
Perfect score: 95
Sequence: 1 VPSYSFIRAHDSFVQDLIA 19

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_41*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	91	95.8	1455	1	GTFC_STRMU
2	91	95.8	1455	1	GTFC_STRMU
3	86	90.5	1597	1	GTFC_STRMU
4	84	88.4	1592	1	GTFC_STRMU
5	73	76.8	1365	1	GTFC_STRMU
6	65	68.4	1462	1	GTFC_STRMU
7	43	45.3	214	1	CMPA_BACST
8	42	44.2	644	1	RNB_ECOLI
9	42	44.2	644	1	RNB_ECOLI
10	42	44.2	659	1	RNB_SHIFL
11	42	44.2	659	1	RNB_SHIFL
12	42	44.2	659	1	RNB_SHIFL
13	41.5	43.7	713	1	DCUZ_ECOLI
14	41	43.2	344	1	MIZD_BACST
15	41	43.2	644	1	RNB_SALTI
16	41	43.2	644	1	RNB_SALTI
17	41	43.2	658	1	RNB_PASMT
18	40	42.1	205	1	ADEN_ADEG8
19	40	42.1	431	1	CRTX_PANAN
20	40	42.1	436	1	THIC_CLOAB
21	40	42.1	548	1	LIP2_CANRU
22	40	42.1	549	1	LIP3_CANRU
23	40	42.1	549	1	LIP4_CANRU
24	40	42.1	572	1	PTUB_STAUV
25	40	42.1	6359	1	BACC_BACLI
26	39.5	41.6	192	1	CUPP_DROPS
27	39	41.1	403	1	COAT_BOOLV
28	39	41.1	509	1	CA43_DROME
29	39	41.1	557	1	WR20_ARATH
30	39	41.1	585	1	ASNS_WAIZE
31	39	41.1	715	1	DOA1_YEAST
32	39	41.1	737	1	SKN1_CANAL
33	39	41.1	946	1	YER6_YEAST

34	39	41.1	1383	1	PRAX_RAT	063425 rattus norv
35	39	41.1	2090	1	N214_HUMAN	P35658 homo sapien
36	39	41.1	2116	1	MY52_DICDI	P08799 dictyosteli
37	38.5	40.5	417	1	PREB_HUMAN	O9HCU5 homo sapien
38	38.5	40.5	454	1	NEM_PIG	P08552 sus scrofa
39	38.5	40.5	499	1	AINX_HUMAN	Q16352 homo sapien
40	38.5	40.5	500	1	PTGI_HUMAN	Q16647 homo sapien
41	38.5	40.5	504	1	AINX_MOUSE	P46660 mus musculu
42	38.5	40.5	505	1	AINX_RAT	P23565 rattus norv
43	38.5	40.5	810	1	NEM_BOVIN	O77788 bos taurus
44	38.5	40.5	845	1	NEM_RAT	P12839 rattus norv
45	38.5	40.5	915	1	NEM_HUMAN	P07197 homo sapien

ALIGNMENTS

RESULT 1
GTFC_STRMU STANDARD: PRT: 1455 AA.
ID GTFC_STRMU
AC P13470; 069382; 069385; 069388; 069391; 069397; P05427;
DT 01-NOV-1988 (Rel. 09, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Glucosyltransferase-SI precursor (EC 2.4.1.5) (GTF-SI)
DE (Dextranucrase) (Sucrose 6-glucosyltransferase).
DN GTF OR SWU 1005
OS Streptococcus mutans.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1309;
RX STRAIN=GS-5;
RX MEDLINE=89137980; PubMed=2976010;
RA Ueda S., Shiroza T., Kuramitsu H.K.;
RT "Sequence analysis of the gtfC gene from Streptococcus mutans GS-5.";
RL Gene 69:101-109(1988).
[2]
RX STRAIN=MT4239 / Serotype C, MT4245 / Serotype E, MT4251 / Serotype F,
RX MT4467 / Serotype E, and MT8148 / Serotype C;
RA MEDLINE=98231643; PubMed=9570124;
RA Fujiwara T., Terao Y., Hoshino T., Kawabata S., Ooshima T., Sobue S.,
RA Kimura S., Hamada S.;
RT "Molecular analyses of glucosyltransferase genes among strains of
RT Streptococcus mutans.";
RL FEMS Microbiol. Lett. 161:331-336(1998).
[3]
RX STRAIN=UA159 / ATCC 700610 / Serotype C;
RX MEDLINE=22295063; PubMed=12397186;
RA Ajdic D., Sherman W.M., McLaughlin R.E., Savic G., Chang J.,
RA Carson M.B., Primateaux C., Tian R., Kenton S., Jia H., Lin S.,
RA Li S., Zhu H., Najjar F., Lai H., White J., Roe B.A., Ferretti J.J.;
RT "Genome sequence of Streptococcus mutans UA159, a cariogenic dental
RT pathogen.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:14434-14439(2002).
[4]
RX STRAIN=GS-5;
RX MEDLINE=87308013; PubMed=3040685;
RA Shiroza T., Ueda S., Kuramitsu H.K.;
RT "Sequence analysis of the gtfB gene from Streptococcus mutans.";
RL J. Bacteriol. 183:4263-4270(1987).
CC -!- FUNCTION: PRODUCTION OF EXTRACELLULAR GLUCANS, THAT ARE THOUGHT
CC TO PLAY A KEY ROLE IN THE DEVELOPMENT OF THE DENTAL PLAQUE BECAUSE
CC OF THEIR ABILITY TO ADHERE TO SMOOTH SURFACES AND MEDIANE THE
CC AGGREGATION OF BACTERIAL CELLS AND FOOD DEBRIS.
CC -!- CATABOLIC ACTIVITY: Sucrose + (1,6)-alpha-D-glucosyl- (N) = D-
CC -!- fructose + (1,6)-alpha-D-glucosyl- (N+1).
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- DISEASE: DENTAL CARIES.

FT	VARIANT	1398	1398	V -> I (IN STRAIN MT8148) .
FT	VARIANT	1424	1424	D -> N (IN STRAIN MT4239) .
FT	VARIANT	1439	1439	V -> I (IN STRAINS MT4239 AND MT8148) .
FT	VARIANT	1444	1444	S -> P (IN STRAIN MT8148) .
FT	CONFLICT	1337	1455	ORLTFKNGVOAKGELLTERGRIRIKYDPNNGSEVNRNVYR TSSGNWYFPGNDGYALIGHVREGRRVYFDENGVRVASHD ORNMMDYDRDDPGRSGSNAVAFPRHSRGSPFNPFRE -> HASTLSLMVPRLRRESSLDOSAVVNSMTMLIPMKRVIVM (IN REF. 1) .
SEQ	SEQUENCE	1455 AA;	162965 MW;	3CB455A99A4FEC86 CRC64;
Qy	Query Match:	95.8%;	Score 91;	DB 1; Length 1455;
Db	Beet Local Similarity	100.0%;	Pred. No. 3.3e-07;	
	Matches 18;	Conservative 0;	Mismatches 0;	Indels 0; Gaps
RESULT 2				
GTFB_STRMU	GTFB_STRMU	STANDARD;	PRT;	1476 AA.
ID	GTFB_SPMU	069387; 069381; 069384; 069387; 069396;		
AC	01-NOV-1988	(Rel. 09, Created)		
DT	18-FEB-2003	(Rel. 41, Last sequence update)		
DT	15-SEP-2003	(Rel. 42, Last annotation update)		
DE	Glucosyltransferase-I precursor (EC 2.4.1.5) (GTF-I) (Dextranucrase)			
DE	(Sucrose 6-glucosyltransferase).			
CN	GTFB OR SMU.1004.			
OC	Streptococcus mutans.			
OC	Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;			
OC	Streptococcus.			
OX	NCB1_Taxid=1309;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN-GS-5;			
RA	MEDLINE=87308013; PubMed=3040685;			
RA	Shiroza T., Ueda S., Kuramitsu H.K.;			
RT	"Sequence analysis of the gtfB gene from Streptococcus mutans."			
RL	J. Bacteriol. 169:4263-4270(1987).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN-MT4239 / Serotype C, MT4245 / Serotype E, MT4251 / Serotype F,			
RC	MT4467 / Serotype E, and MT8148 / Serotype C;			
RC	MEDLINE=98231643; PubMed=9570124;			
RA	Fujisawa T., Terao Y., Hoshino T., Kawabata S., Sobue S.,			
RA	Kimura S., Hamada S.;			
RT	"Molecular analyses of glucosyltransferase genes among strains of			
RT	Streptococcus mutans."			
RL	FEMS Microbiol. Lett. 161:331-336(1998).			
RN	[3]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN-UAI59 / ATCC 700610 / Serotype C;			
RA	MEDLINE=22295063; PubMed=12397186;			
RA	Ajdle D., Woshan W.M., McLaughlin R.E., Savic G., Chang J.,			
RA	Carson M.B., Primeaux C., Tian R., Kenton S., Via H., Lin S., Qian Y.,			
RA	Li S., Zhu H., Najaf F., Lai H., White J., Roe B.A., Perretti J.D.;			
RT	"Genome sequence of Streptococcus mutans UAI59, a cariogenic dental			
RT	pathogen."			
RL	Proc. Natl. Acad. Sci. U.S.A. 99:14434-14439(2002).			
CC	-1- FUNCTION: PRODUCTION OF EXTRACELLULAR GLUCANS, THAT ARE THOUGHT			
CC	TO PLAY A KEY ROLE IN THE DEVELOPMENT OF THE DENTAL PLAQUE BECAUSE			
CC	OF THEIR ABILITY TO ADHERE TO SMOOTH SURFACES AND MEDiate THE			
CC	AGGREGATION OF BACTERIAL CELLS AND FOOD DEBRIS.			
CC	-1- CATALYTIC ACTIVITY: Sucrose + {(1,6)-alpha-D-glucosyl} (N)= D-			
CC	fructose + {(1,6)-alpha-D-glucose}			
CC	-1- SUBCELLULAR LOCATION: Secreted.			
CC	-1- DISEASE: DENTAL CARIES.			
CC	-1- MISCELLANEOUS: GTF-I SYNTHESIZES WATER-INSOLUBLE GLUCANS (ALPHA			
CC	1,3-LINKED GLUCOSE AND SOME 1,6 LINKAGES), GTF-S SYNTHESIZES			
CC	WATER-SOLUBLE GLUCANS (ALPHA 1,6-GLUCOSE). GTF-SI SYNTHESIZES BOTH			
CC	FORMS OF GLUCANS.			

CC	-1-	SIMILARITY:	TO OTHER GLUCOSYLTRANSFERASES AND SOME TO A GLUCAN-BINDING PROTEIN FROM S. MUTANS.
CC	-1-	SIMILARITY:	Contains 10 cell wall binding repeats.
CC	-----		
CC	CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).	
CC	-----		
DR	EMBL;	M17361;	AAA85858.1; ; -
DR	EMBL;	D86651;	BAA26101.1; ; -
DR	EMBL;	D86654;	BAA26105.1; ; -
DR	EMBL;	D86657;	BAA26109.1; ; -
DR	EMBL;	D86660;	BAA26113.1; ; -
DR	EMBL;	D89977;	BAA26119.1; ; -
DR	EMBL;	AEO14940;	AAN58705.1; ; -
DR	InterPro;	IPR002479;	CW binding.
DR	InterPro;	IPR00318;	Glyco hydro_70.
DR	Pfam;	PF01473;	CW binding_I; 1..1.
DR	Pfam;	PF02324;	Glyco hydro_70; 1.
KM	Transferase;	Glycosyltransferase;	Signal; Repeat; Dental carrier;
KM	Complete proteome.		
FT	SIGNAL	1	34
FT	CHAIN	35	1476
FT	DOMAIN	35	1051
FT	DOMAIN	1097	1476
FT	REPEAT	1097	1130
FT	DOMAIN	1161	1470
FT	REPEAT	1161	1210
FT	REPEAT	1225	1275
FT	REPEAT	1290	1340
FT	REPEAT	1355	1405
FT	REPEAT	1420	1470
FT	VARIANT	62	62
FT	VARIANT	65	65
FT	VARIANT	68	68
FT	VARIANT	68	68
FT	VARIANT	78	78
FT	VARIANT	86	86
FT	VARIANT	89	89
FT	VARIANT	168	168
FT	VARIANT	276	276
FT	VARIANT	399	399
FT	VARIANT	474	474
FT	VARIANT	512	512
FT	VARIANT	519	519
FT	VARIANT	701	701
FT	VARIANT	708	708
FT	VARIANT	938	938
FT	VARIANT	952	957
FT	VARIANT	963	964
FT	VARIANT	968	970
FT	VARIANT	1086	1086
FT	VARIANT	1158	1158
FT	VARIANT	1163	1163
FT	VARIANT	1168	1168
FT	VARIANT	1182	1182
FT	VARIANT	1234	1234
FT	VARIANT	1263	1263
FT	VARIANT	1264	1264
FT	VARIANT	1272	1272
FT	VARIANT	1329	1329

```

FT  VARIANT 1394 1394 Y -> H (IN STRAINS GS-5, MT4239, MT4467
FT  VARIANT 1402 1402 S -> G (IN STRAINS GS-5, MT4239, MT4467
FT  VARIANT 1459 1459 AND MT8148).
FT  VARIANT 1459 1459 Y -> H (IN STRAIN MT4467).
FT  CONFLICT 570 570 R -> A (IN REF. 1).
FT  CONFLICT 800 817 ADDOVERVASTASTASTDCK -> LKFMALMLAPBHQWA
FT  CONFLICT 1310 1310 (IN REF. 1).
FT  CONFLICT 1476 AA; 165846 MM; 9CCE03F731B4CBFC CRC64;
SQ  SEQUENCE

Query Match 95.8%; Score 91; DB 1; Length 1476;
Best Local Similarity 100.0%; Pred. No. 3.3e-07;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 VPSSYFIRAHDSYVDLI 18
Db 552 VPSSYFIRAHDSYVDLI 569

RESULT 3
GTF1_STRDO STANDARD; PRT; 1597 AA.
ID_GTF1_STRDO
AC p11001;
DT 01-JUN-1989 (Rel. 11, Created)
DT 01-JUN-1989 (Rel. 11, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Glucosyltransferase-I precursor (EC 2.4.1.5) (GTF-I) (Dextranucrase)
DE (Sucrose 6-glucosyltransferase).
GN GTFI.
OS Streptococcus downei (Streptococcus sobrinus).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=13117;
RX MEDLINE=87308014; PubMed=3040686;
RA Ferretti J.J., Gilpin M.L., Ruseell R.R.B.;
RT "Nucleotide sequence of a glucosyltransferase gene from Streptococcus
RT sobrinus MFe28."
PL J. Bacteriol. 169:4271-4278 (1987).
CC -1- FUNCTION: PRODUCTION OF EXTRACELLULAR GLUCANS. THAT ARE THOUGHT
CC TO PLAY A KEY ROLE IN THE DEVELOPMENT OF THE DENTAL PLAQUE BECAUSE
CC OF THEIR ABILITY TO ADHERE TO SMOOTH SURFACES AND MEDIATE THE
CC AGGREGATION OF BACTERIAL CELLS AND FOOD DEBRIS.
CC -1- CATALYTIC ACTIVITY: Sucrose + ((1,6)-alpha-D-glucosyl) (N) = D-
CC fructose + ((1,6)-alpha-D-glucosyl) (N+1).
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- DISEASE: DENTAL CARIES.
CC -1- MISCELLANEOUS: GTF-I SYNTHESIZES WATER-INSOLUBLE GLUCANS (ALPHA
CC 1,3-LINKED GLUCOSE AND SOME 1,6 LINKAGS), GTF-S SYNTHESIZES
CC WATER-SOLUBLE GLUCANS (ALPHA 1,6-GLUCOSE). GTF-S1 SYNTHESIZES BOTH
CC FORMS OF GLUCANS.
CC -1- SIMILARITY: TO OTHER GLUCOSYLTRANSFERASES AND SOME TO A GLUCAN-
CC BINDING PROTEIN FROM S. MUTANS.
CC -1- SIMILARITY: Contains 19 cell wall binding repeats.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M17391; AAC63063.1;
CC InterPro; IPR002479; CW binding.
CC InterPro; IPR003318; Glyco_hydro_70.
CC Pfam; PF01473; CW binding_1; 16.
CC Pfam; PF02324; Glyco_hydro_70; 1.
CC Transferrase; Glycosyltransferase; Signal; Repeat; Dental caries.
FT 1 38 POTENTIAL.

```

```

FT CHAIN 39 1597 GLUCOSYLTRANSFERASE-I.
FT DOMAIN 39 1050 CATALYTIC (APPROXIMATE).
FT DOMAIN 1099 1597 GLUCAN-BINDING (APPROXIMATE).
FT DOMAIN 1099 1597 1.25 A, 2 B, AND 5 AC REPEATS.
FT REPEAT 1099 1597 A REPEAT.
FT REPEAT 1163 1213 AC REPEAT.
FT REPEAT 1227 1277 AC REPEAT.
FT REPEAT 1292 1342 AC REPEAT.
FT REPEAT 1352 1399 B REPEAT.
FT REPEAT 1406 1455 AC REPEAT.
FT REPEAT 1465 1512 AC REPEAT.
FT REPEAT 1519 1568 AC REPEAT.
FT REPEAT 1582 1597 A REPEAT (INCOMPLETE).
SQ SEQUENCE 1597 AA; 177080 MM; B9E86A2008687998 CRC64;

Query Match 90.5%; Score 86; DB 1; Length 1597;
Best Local Similarity 94.4%; Pred. No. 2.5e-06;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 VPSYSTRANDSEVODLI 18
Db 554 VPSYSTRANDSEVODLI 571

RESULT 4
GTF2_STRDO STANDARD; PRT; 1592 AA.
ID GTF2_STRDO
AC P27470;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Glucosyltransferase-I precursor (EC 2.4.1.5) (GTF-I) (Dextranucrase)
DE (Sucrose 6-glucosyltransferase).
OS Streptococcus downei (Streptococcus sobrinus).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OC NCBI_TaxID=1317;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=6715 / Serotype G;
RX MEDLINE=91123227; PubMed=1704006;
RA Abo H., Matsumura T., Kodama T., Ohba H., Kato K.,
RA Kagawa H.;
RT "peptide sequences for sucrose splitting and glucan binding within
RT Streptococcus sobrinus glucosyltransferase (water-insoluble glucan
RT synthetase)"; 173:989-996(1991).
RL J. Bacteriol. 173:989-996(1991).
CC -1- FUNCTION: PRODUCTION OF EXTRACELLULAR GLUCANS, THAT ARE THOUGHT
CC TO PLAY A KEY ROLE IN THE DEVELOPMENT OF THE DENTAL PLAQUE BECAUSE
CC OF THEIR ABILITY TO ADHERE TO SMOOTH SURFACES AND MEDIATE THE
CC AGGREGATION OF BACTERIAL CELLS AND FOOD DEBRIS.
CC -1- CATALYTIC ACTIVITY: Sucrose + ((1,6)-alpha-D-glucosyl) (N) = D-
CC fructose + ((1,6)-alpha-D-glucosyl) (N+1).
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- DISEASE: DENTAL CARIES.
CC -1- MISCELLANEOUS: GTF-I SYNTHESIZES WATER-INSOLUBLE GLUCANS (ALPHA
CC 1,3-LINKED GLUCOSE AND SOME 1,6 LINKAGES), GTF-S SYNTHESIZES
CC WATER-SOLUBLE GLUCANS (ALPHA 1,6-GLUCOSE). GTF-SI SYNTHESIZES BOTH
CC FORMS OF GLUCANS.
CC -1- SIMILARITY: TO OTHER GLUCOSYLTRANSFERASES AND SOME TO A GLUCAN-
CC BINDING PROTEIN FROM S.MUTANS.
CC -1- SIMILARITY: Contains 16 cell wall binding repeats.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; D90213; BAAL4241.1; -
CC InterPro; IPR002479; CW_binding.

```

```

DR InterPro; IPR003318; Glyco_hydro_70.
DR Pfam; PF01473; CW_binding_1; 13.
DR Pfam; PF02324; Glyco_hydro_70; 1.
KM Transferase; Glycosyltransferase; Signal; Repeat; Dental caries.
FT CHAIN 1 38 POTENTIAL.
FT CHAIN 39 1592 GLUCOSYLTRANSFERASE-I.
FT DOMAIN 39 1044 CATALYTIC (APPROXIMATE).
FT DOMAIN 1093 1592 GLUCAN-BINDING (APPROXIMATE).
FT DOMAIN 1093 1592 6-5 X TANDEM REPEATS.
FT REPEAT 1093 1142 1.
FT REPEAT 1158 1207 2.
FT REPEAT 1222 1272 3.
FT REPEAT 1287 1337 4.
FT REPEAT 1402 1451 5.
FT REPEAT 1514 1563 6.
FT REPEAT 1577 1592 7 (INCOMPLETE).
SQ SEQUENCE 1592 AA; 176167 MM; BC0A6D079351ECF CRC64;

Query Match 88.4%; Score 84; DB 1; Length 1592;
Best Local Similarity 88.9%; Pred. No. 5.5e-06;
Matches 16; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 VPSYSTRANDSEVODLI 18
Db 548 VPSYSTRANDSEVODLI 565

RESULT 5
GTF5_STRDO STANDARD; PRT; 1365 AA.
ID GTF5_STRDO
AC P29356;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Glucosyltransferase-S precursor (EC 2.4.1.5) (GTF-S) (Dextranucrase)
DE (Sucrose 6-glucosyltransferase).
GN GTF5.
OS Streptococcus downei (Streptococcus sobrinus).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OC NCBI_TaxID=1317;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MFE28;
RX MEDLINE=90316665; PubMed=2142479;
RA Gilmore K.S., Russell J.R., Ferretti J.J.;
RT "analysis of the Streptococcus downei gtfS gene, which specifies a
RT glucosyltransferase that synthesizes soluble glucans";
RL Infect. Immun. 58:2452-2458(1990).
CC -1- FUNCTION: PRODUCTION OF EXTRACELLULAR GLUCANS, THAT ARE THOUGHT
CC TO PLAY A KEY ROLE IN THE DEVELOPMENT OF THE DENTAL PLAQUE BECAUSE
CC OF THEIR ABILITY TO ADHERE TO SMOOTH SURFACES AND MEDIATE THE
CC AGGREGATION OF BACTERIAL CELLS AND FOOD DEBRIS.
CC -1- CATALYTIC ACTIVITY: Sucrose + ((1,6)-alpha-D-glucosyl) (N) = D-
CC fructose + ((1,6)-alpha-D-glucosyl) (N+1).
CC -1- ENZYME REGULATION: GLUCAN SYNTHESIS BY GTF-S IS INDEPENDENT OF
CC PRIMER GLUCAN UNLIKE GTF-I.
CC -1- DISEASE: DENTAL CARIES.
CC -1- MISCELLANEOUS: GTF-S SYNTHESIZES WATER-SOLUBLE GLUCANS (ALPHA
CC 1,6-GLUCOSE).
CC -1- SIMILARITY: TO OTHER GLUCOSYLTRANSFERASES AND SOME TO A GLUCAN-
CC BINDING PROTEIN FROM S.MUTANS.
CC -1- SIMILARITY: Contains 10 cell wall binding repeats.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M30943; AAA26898.1; -

```

```

DR InterPro; IPR002479; CM_binding.
DR InterPro; IPR003318; Glyco_hydro_70.
DR Pfam; PF01473; CM_binding_1; 8.
DR Pfam; PF02324; Glyco_hydro_70; 1.
KW Transferrase; Glycosyltransferase; Signal; Repeat; Dental caries.
FT SIGNAL 1 36 OR 37 (POTENTIAL).
FT CHAIN 37 1365 GLUCOSYLTRANSFERASE-S.
FT REPEAT 157 177 CELL WALL BINDING 1.
FT REPEAT 178 197 CELL WALL BINDING 2.
FT DOMAIN 198 1061 CATALYTIC (APPROXIMATE).
FT REPEAT 1062 1082 CELL WALL BINDING 3.
FT REPEAT 1083 1102 CELL WALL BINDING 4.
FT REPEAT 1150 1169 CELL WALL BINDING 5.
FT REPEAT 1170 1190 CELL WALL BINDING 6.
FT REPEAT 1225 1243 CELL WALL BINDING 7.
FT REPEAT 1289 1308 CELL WALL BINDING 8.
FT REPEAT 1309 1328 CELL WALL BINDING 9.
FT REPEAT 1331 1352 CELL WALL BINDING 10.
SQ SEQUENCE 1365 AA; 151590 MW; 1672965A2E8C476 CRC64;

Query Match 76.8%; Score 73; DB 1; Length 1365;
Best Local Similarity 78.9%; Pred. No. 0.00034;
Matches 15; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 1 VPSYSEFIRAHDSVODLTA 19
Db 537 VPSYSEFIRAHDSVODLTA 555

RESULT 6
GTFD_STRMU STANDARD; PRT; 1462 AA.
ID GTFD_STRMU STANDARD; PRT; 1462 AA.
AC P43331; O69383; O69386; O69389; O69392; O69398;
DT 01-FEB-1996 (Rel. 33, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Glucosyltransferase-S precursor (EC 2.4.1.5) (GTF-S) (Dextranucrase)
DE (Sucrose 6-glucosyltransferase).
GN GTFD OR SMU.910.
OS Streptococcus mutans.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OC NCB1_taxid=1309;
CX [1]
RP SEQUENCE FROM N.A.
RC STRAIN=GS-5;
RX MEDLINE=91100958; PubMed=2148600;
RA Honda O., Kato C., Kuramitsu H.K.;
RT "Nucleotide sequence of the Streptococcus mutans gtfD gene encoding
the glucosyltransferase-S enzyme."
RL J. Gen. Microbiol. 136:2099-2105(1990).
[2]
RP SEQUENCE FROM N.A.
RC STRAIN=MT4239 / Serotype C, MT4245 / Serotype E, MT4251 / Serotype F,
MT4467 / Serotype B, and MT8148 / Serotype C;
RX MEDLINE=98231643; PubMed=9570124;
RA Fujiwara T., Terao Y., Hoshino T., Kawabata S., Ooshima T., Sobue S.,
RA Kimura S., Hamada S.;
RT "Molecular analyses of glucosyltransferase genes among strains of
Streptococcus mutans."
RL PEMS Microbiol. Lett. 161:331-336(1998).
[3]
RP SEQUENCE FROM N.A.
RC STRAIN=UA159 / ATCC 700610 / Serotype C;
RX MEDLINE=22295063; PubMed=12397186;
RA Ajdic D., McShan W.M., McLaughlin R.B., Savic G., Chang J.,
RA Carson M.B., Primeaux C., Tian R., Kenton S., Jia H., Ilin S., Qian Y.,
RA Li S., Zhu H., Najjar F., Lai H., White J., Roe B.A., Ferretti J.J.;
RT "Genome sequence of Streptococcus mutans UA159, a cariogenic dental
pathogen."
RL Proc. Natl. Acad. Sci. U.S.A. 99:14434-14439(2002).
CC -1- FUNCTION: PRODUCTION OF EXTRACELLULAR GLUCANS, THAT ARE THOUGHT
TO PLAY A KEY ROLE IN THE DEVELOPMENT OF THE DENTAL PLAQUE BECAUSE

```

```

CC OF THEIR ABILITY TO ADHERE TO SMOOTH SURFACES AND MEDIANE THE
CC AGGREGATION OF BACTERIAL CELLS AND FOOD DEBRIS.
CC -1- CATALYTIC ACTIVITY: Sucrose + ((1,6)-alpha-D-glucosyl) (N) = D-
CC fructose + ((1,6)-alpha-D-glucosyl) (N+1).
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- DISEASE: DENTAL CARIES.
CC -1- MISCELLANEOUS: GTF-I SYNTHESIZES WATER-INSOLUBLE GLUCANS (ALPHA
CC 1,3-LINKED GLUCOSE AND SOME 1,6 LINKAGES), GTF-S SYNTHESIZES
CC WATER-SOLUBLE GLUCANS (ALPHA 1,6-GLUCOSE), GTF-SI SYNTHESIZES BOTH
CC FORMS OF GLUCANS.
CC -1- SIMILARITY: TO OTHER GLUCOSYLTRANSFERASES AND SOME TO A GLUCAN-
CC BINDING PROTEIN FROM S. MUTANS.
CC -1- SIMILARITY: Contains 6 cell wall binding repeats.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.ebi.ac.uk/announcements/
CC or send an email to license@ebi.ac.uk).
CC -----
CC EMBL; M29296; AAA26895.1; -
CC EMBL; D88653; BAA26103.1; -
CC EMBL; D88656; BAA26107.1; -
CC EMBL; D88659; BAA26111.1; -
CC EMBL; D88662; BAA26115.1; -
CC EMBL; D89979; BAA26121.1; -
CC EMBL; AE014932; AAN58619.1; -
CC InterPro; IPR002479; CM_binding.
CC InterPro; IPR003318; Glyco_hydro_70.
DR Pfam; PF02324; Glyco_hydro_70.1.
KW Transferrase; Glycosyltransferase; Signal; Repeat; Dental caries;
KW Complete proteome.
FT CHAIN 1 1462 ?
FT DOMAIN 1232 1423 1462
FT REPEAT 1232 1295
FT REPEAT 1296 1359
FT REPEAT 1360 1423
FT VARIANT 10 10
FT 19 19
FT VARIANT 58 58
FT 68 68
FT VARIANT 81 81
FT 113 113
FT VARIANT 122 122
FT 132 132
FT VARIANT 135 135
FT 137 137
FT VARIANT 202 202
FT 255 255
FT VARIANT 275 275
FT 288 288
FT VARIANT 301 301
FT 313 313
FT VARIANT 317 317
FT 328 328
FT VARIANT 350 350
FT 628 633
FT 688 688
FT 726 732
FT 726 730
FT 762 762
FT VARIANT 762 762

```

```

FT  VARIANT  964      964      D -> Y (IN STRAIN MT4251).
FT  VARIANT  1019     1019     E -> K (IN STRAINS MT4245 AND MT4251).
FT  VARIANT  1059     1059     LG -> IR (IN STRAIN MT4251).
FT  VARIANT  1060     1060     G -> R (IN STRAIN MT4245).
FT  VARIANT  1080     1080     G -> R (IN STRAIN MT4245).
FT  VARIANT  1142     1142     H -> Q (IN STRAIN GS-5).
FT  VARIANT  1198     1198     S -> N (IN STRAIN MT4239).
FT  VARIANT  1220     1220     Y -> C (IN STRAINS MT4251 AND MT4467).
FT  VARIANT  1280     1280     F -> L (IN STRAIN MT4467).
FT  VARIANT  1282     1282     Q -> P (IN STRAIN MT4245).
FT  VARIANT  1290     1290     K -> T (IN STRAIN MT4245).
FT  VARIANT  1311     1311     N -> D (IN STRAIN MT4245).
FT  VARIANT  1403     1403     G -> D (IN STRAIN GS-5).
FT  VARIANT  1425     1425     G -> R (IN STRAIN GS-5).
FT  VARIANT  1449     1449     R -> K (IN STRAIN MT4467).
FT  CONFLICT  1448     1448     RYDDKSGMMVKNKVVTLNNGRRIGIDRWGIARY -> VY
FT  CONFLICT  1462     1462     R (IN REF. 1).
SQ  SEQUENCE  1462 AA;  163387 MW;  CEAA279CAD708645 CRC64;

```

```

Query Match      68.4%; Score 65; DB 1; Length 1462;
Best Local Similarity 76.5%; Pred. No. 0.0085;
Matches 13; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

```

```

OY  3 SYSFIRAHDSVQDLIA 19
    :|||||:
Db  576 NYIFRAHDSVQTVIA 592

```

```

RESULT 7
CMPA_BACSU
ID  _CMPA_BACSU  STANDARD; PRT; 214 AA.
AC  P14264;
DT  01-JAN-1990 (Rel. 13, Created)
DT  01-JAN-1990 (Rel. 13, Last sequence update)
DT  15-SEP-2003 (Rel. 42, Last annotation update)
DE  Competence protein A.
GN  COMA OR COMA OR COMA.
OS  Bacillus subtilis.
OC  Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX  NCBI_TaxID=1423;
RN  [1]
RP  SEQUENCE FROM N.A.
RX  MEDLINE=90008771; PubMed=2507523;
RT  Weinrauch Y., Guillen N., Dubnau D.;
RT  "Sequence and transcription mapping of Bacillus subtilis competence
RT  genes comB and comA, one of which is related to a family of bacterial
RT  regulatory determinants."
RL  J. Bacteriol. 171:5362-5375(1989).
RN  [2]
RP  SEQUENCE FROM N.A.
RX  MEDLINE=97419515; PubMed=9274030;
RA  Oudega B., Koningsstein G., Rodriguez L., de Sales Ramon M.,
RA  Hilbert H., Duesterhoeft A., Pohl T.M., Weitzenecker T.;
RT  "Analysis of the Bacillus subtilis genome: cloning and nucleotide
RT  sequence of a 62 kb region between 275 degrees (rrnB) and 284 degrees
RT  (gal).";
RL  Microbiology 143:2769-2774(1997).
RN  [3]
RP  SEQUENCE FROM N.A.
RX  MEDLINE=98044033; PubMed=9384377;
RA  Kunst F., Ogasawara N., Noszer I., Albertini A.M., Alloni G.,
RA  Azevedo V., Beterlo M.G., Bessieres P., Bolotin A., Borchert S.,
RA  Boursier R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,
RA  Brouillet S., Brunsch C.V., Caldwell B., Capuano V., Carter N.M.,
RA  Choi S.K., Codani J.J., Comerston I.F., Cummings N.J., Daniel R.A.,
RA  Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emerson P.T.,
RA  Entian K.D., Ertugron J., Fabre C., Ferrari B., Foulger D.,
RA  Fitz C., Fujita M., Fujita Y., Fuma S., Galazzi A., Galleton N.,
RA  Ghit S.Y., Glaser P., Goffeau A., Goldightly E.J., Grand G.,
RA  Guiseppi G., Guy B.J., Haeck U., Harwood C.R., Henaut A.,
RA  Hilbert H., Holappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,

```

```

RA  Joris B., Karamata D., Kasahara Y., Klaerr-Blanchard M., Klein C.,
RA  Kobayashi Y., Koetter P., Koningsstein G., Krogh S., Kumano M.,
RA  Kurita K., Lapidus A., Lardinois S., Lander J., Lazarevic V.,
RA  Lee S.M., Levine A., Liu H., Masuda S., Mauel C., Medigic C.,
RA  Medina N., Mellado R.F., Mizuno M., Moesti D., Nakai S., Noback M.,
RA  Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,
RA  Paro V., Pohl T.M., Portelle D., Porwollik S., Prescott A.M.,
RA  Priescan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,
RA  Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadele Y.,
RA  Sato T., Scanlan E., Schleicher S., Schroeder R., Scoffone F.,
RA  Sekiguchi U., Sekowska A., Seror S.U., Serro P., Shin B.S., Solde B.,
RA  Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,
RA  Takeuchi M., Tamakoshi A., Tanaka T., Terpestra P., Tognoni A.,
RA  Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,
RA  Viari A., Wandut R., Wedler E., Wedler H., Weitzenecker T.,
RA  Winers P., Wipat A., Yamamoto H., Yamane K., Yasunoto K., Yata K.,
RA  Yoshida K., Yoshikawa H.F., Zumbstein E., Yoshikawa H., Danchin A.;
RT  "The complete genome sequence of the Gram-positive bacterium Bacillus
RT  subtilis.";
RL  Nature 390:249-256(1997).
CC  -!- FUNCTION: Required for the development of competence in
CC  B.subtilis, for the expression of late-expressing competence
CC  genes, and for the expression of the growth stage-regulated
CC  molecule surfactin. May play a regulatory role during the
CC  development of competence, and an analogous role as uvrC-ORF2
CC  (transducing environmental information to the DNA repair system).
CC  -!- PTM: Phosphorylated by comB.
CC  -!- SIMILARITY: Contains 1 response regulatory domain.
CC  -!- SIMILARITY: BELONGS TO THE LUXR/UHPA FAMILY OF TRANSCRIPTIONAL
CC  REGULATORS.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@sib-sib.ch).
CC  -----
DR  EMBL, M22856; AAA22320.1; -
DR  EMBL, Z93932; CAB07904.1; -
DR  EMBL, Z99120; CAB15156.1; -
DR  PIR, A31591; RGSBCA.
DR  Subtilist; BG10381; comA.
DR  InterPro: IPR000792; HTH_LuxR.
DR  InterPro: IPR001789; Response_reg.
DR  Pfam: PF00196; GcrE; 1.
DR  Pfam: PF00072; response_reg; 1.
DR  PRINTS; PR00038; HTH_LuxR.
DR  ProDom; PD000307; HTH_LuxR; 1.
DR  ProDom; PD000039; Response_reg; 1.
DR  SMART; SM00421; HTH_LuxR; 1.
DR  SMART; SM00448; REC; 1.
DR  PROSITE; PS00622; HTH_LuxR_FAMILY; 1.
DR  PROSITE; PS0110; RESPONSE_REGULATORY; 1.
KM  Competence; Sensory transduction; Transcription regulation; Activator;
KW  DNA-binding; Phosphorylation; Complete proteome.
FT  DOMAIN 1 121 RESPONSE REGULATORY.
FT  MOD_RES 55 55 PHOSPHORYLATION (BY SIMILARITY).
FT  DNA_BIND 171 190 H-T-H MOTIF (BY SIMILARITY).
SQ  SEQUENCE 214 AA; 24128 MW; 52D01A9740759072 CRC64;

```

```

Query Match      45.3%; Score 43; DB 1; Length 214;
Best Local Similarity 52.9%; Pred. No. 5.4;
Matches 9; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

```

```

OY  2 PSYSFIRAHDSVQDLI 18
    :|||||:
Db  36 PSEGFIRKQHDFFSSYDLI 52

```

```

RESULT 8
RNB_ECO57

```


DR EMBL: D90766; BAA14840.1; -.
DR F1R; A64877; A64877.
DR EcoGene; Egl1620; rnb.
DR HAMAP; MF_01036; -; 1.
DR InterPro; IPR004476; 3_prime_RNase.
DR InterPro; IPR002059; Cold_shock.
DR InterPro; IPR001900; Ribonuclease_II.
DR InterPro; IPR003029; SL.
DR Pfam; PF00773; RNB; 1.
DR SMART; SM00357; CSP; 1.
DR SMART; SM00316; S1; 1.
DR TIGRfam; TIGR00358; 3_prime_RNase; 1.
DR PROSITE; PS01175; RIBONUCLEASE_II; 1.
DR PROSITE; PS50126; SL; FALSE NEG.
DR HydroLase; Exonuclease; Nuclease; RNA-binding; Complete proteome.
DR DOMAIN 561 643 S1 MOTIF.
FT CONFLICT 384 384 I -> N (IN REF. 1).
FT CONFLICT 399 399 C -> G (IN REF. 1).
FT CONFLICT 513 513 A -> R (IN REF. 1).
SQ SEQUENCE 644 AA; 72490 MW; 36B16712CDF14394 CRC64;

Query Match 44.2%; Score 42; DB 1; Length 644;
Best Local Similarity 44.4%; Pred. No. 27;
Matches 8; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

QY 2 PSYSTRANDSEVODLIA 19
DB 362 PDYRFLGKGEVLDIVA 379

RESULT 10
RNB_SHIFL ID RNB_SHIFL STANDARD; PRT; 644 AA.
AC P59107;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Exoribonuclease II (EC 3.1.13.1) (Ribonuclease II) (RNase II).
GN RNB OR SFI291.
OS Shigella flexneri.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Shigella.
OX NCBI_TaxID=623;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=301 / Serotype 2a;
RX MEDLINE=22272406; PubMed=12384590;
RA Jin Q., Yuan Z., Xu J., Wang Y., Shen Y., Lu W., Wang J., Liu H.,
RA Sun L., Xue Y., Zhang X., Zhang J., Yang G., Wu H., Qu D., Dong J.,
RA Sun L., Xue Y., Zhao Y., Gao Y., Zhu J., Kan B., Ding K., Chen S.,
RA Cheng H., Yao Z., He B., Chen R., Ma D., Qiang B., Wen Y., Hou Y.,
RA Yu J.;
RT "Genome sequence of Shigella flexneri 2a: insights into pathogenicity
RT through comparison with genomes of Escherichia coli K12 and O157.";
RL Nucleic Acids Res. 30:4432-4441(2002).
CC -1- FUNCTION: Involved in mRNA degradation. Hydrolyzes single-stranded
CC polyribonucleotides processively in the 3' to 5' direction (By
CC similarity).
CC -1- CATALYTIC ACTIVITY: Exonucleolytic cleavage in the 3'- to 5'-
CC direction to yield nucleoside 5'-phosphates.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: Belongs to the ribonuclease II (RNB) family.
CC -1- SIMILARITY: Contains 1 S1 motif domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL: AE015155; AAN42902.1; -.
DR HAMAP; MF_01036; -; 1.
DR InterPro; IPR004476; 3_prime_RNase.
DR InterPro; IPR002059; Cold_shock.
DR InterPro; IPR001900; Ribonuclease_II.
DR InterPro; IPR003029; SL.
DR Pfam; PF00773; RNB; 1.
DR SMART; SM00357; CSP; 1.
DR SMART; SM00316; S1; 1.
DR TIGRfam; TIGR00358; 3_prime_RNase; 1.
DR PROSITE; PS01175; RIBONUCLEASE_II; 1.
DR PROSITE; PS50126; SL; FALSE NEG.
DR HydroLase; Exonuclease; Nuclease; RNA-binding.
DR DOMAIN 561 643 S1 MOTIF.
SQ SEQUENCE 644 AA; 72464 MW; 137E8124EF3735C2 CRC64;

Query Match 44.2%; Score 42; DB 1; Length 644;
Best Local Similarity 44.4%; Pred. No. 27;
Matches 8; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

QY 2 PSYSTRANDSEVODLIA 19
DB 362 PDYRFLGKGEVLDIVA 379

RESULT 11
RNB_HAEIN ID RNB_HAEIN STANDARD; PRT; 659 AA.
AC P44440;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Exoribonuclease II (EC 3.1.13.1) (Ribonuclease II) (RNase II).
GN RNB OR H11733.
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC Pasteurellaceae; Haemophilus.
OX NCBI_TaxID=727;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Rd / KW20 / ATCC 51907;
RX MEDLINE=95350630; PubMed=7542800;
RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
RA Kerlavage A.R., Bult C.J., Tomb J.F., Dougherty B.A., Merrick J.M.,
RA McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
RA Scott J.D., Shirley R., Liu L.-L., Glodek A., Kelley J.M.,
RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
RA Uettermann T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,
RA Fine L.D., Fritchman J.L., Fuhrmann J.L., Geoghagen N.S.M.,
RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
RA Venter J.C.;
RT "Whole-genome random sequencing and assembly of Haemophilus influenzae
RT Rd.";
RL Science 269:496-512(1995).
CC -1- FUNCTION: Involved in mRNA degradation. Hydrolyzes single-stranded
CC polyribonucleotides processively in the 3' to 5' direction (By
CC similarity).
CC -1- CATALYTIC ACTIVITY: Exonucleolytic cleavage in the 3'- to 5'-
CC direction to yield nucleoside 5'-phosphates.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: Belongs to the ribonuclease II (RNB) family.
CC -1- SIMILARITY: Contains 1 S1 motif domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL: U32846; AAC23378.1; -.

```

DR PIR; A64139; A64139.
DR TIGR; H11733; -.
DR HAMAP; MF_01036; -.
DR InterPro; IPR004476; 3_prime_RNase.
DR InterPro; IPR002059; Cold_shock.
DR InterPro; IPR001900; Ribonuclease_II.
DR InterPro; IPR003029; S1.
DR Pfam; PF00773; RNb; 1.
DR Pfam; PF00575; S1; 1.
DR SMART; SM00357; CSP; 1.
DR SMART; SM00316; S1; 1.
DR TIGRfams; TIGR00358; 3_prime_RNase; 1.
DR PROSITE; PS01175; RIBONUCLEASE_II; 1.
DR PROSITE; PS50126; S1; 1.
DR Hydrolase; Exonuclease; Nuclease; RNA-binding; Complete proteome.
KW DOMAIN 576 658 S1 MOTIF.
SQ SEQUENCE 659 AA; 75782 MW; FDDDA9097A7D3B66 CRC64;

Query Match 44.2%; Score 42; DB 1; Length 659;
Best Local Similarity 38.9%; Pred. No. 27;
Matches 7; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

OY 2 PSYSTRANDSEVODLIA 19
DB 367 PDYAFVLAENGKVOEIKK 384

RESULT 12
HTRL SCHPO STANDARD; PRT; 932 AA.
ID HTRL SCHPO
AC P87314;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Histone transcription regulator 1 homolog.
GN SPBC31F10.13C.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RX MEDLINE=21848401; PubMed=11859360;
RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
RA Sgourou J., Peat N., Hayles J., Baker S., Bashan D., Bowman S.,
RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA Collins M., Connor R., Cronin A., Davis P., Felwell T., Fraser A.,
RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,
RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
RA Skelton J., Simmonds M., Squares R., Stevens S., Warren T., Whitehead S.,
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA Woodward J., Volkart G., Aert R., Robben J., Gymoniere B.,
RA Wellens I., Vanterre E., Rieger M., Schaefer M., Meller-Auer S.,
RA Gabel C., Fuchs M., Fritze C., Holzer E., Koeck D., Hilbert H.,
RA Borzym K., Langer M., Beck A., Leinach H., Reinhardt R., Pohl T.M.,
RA Eger P., Zimmermann W., Medler H., Wambutt R., Purnelle B.,
RA Goffeau A., Cadieu F., Dreano S., Gloux S., Lelaure V., Mottier S.,
RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RA Lucas M., Rochet M., Gallardin C., Tallada V.A., Garçon A., Thode G.,
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forburg S.L.,
RA Cerutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
RA Shpkovski G.V., Ussery D., Barrett B.G., Nurse P.;
RT "The genome sequence of Schizosaccharomyces pombe";
RL Nature 415:871-880(2002).
CC -!- FUNCTION: COULD PLAY A PART IN MECHANISMS OF TRANSCRIPTIONAL
REGULATION SIMILAR TO THAT PLAYED BY YEAST HIR1 AND HIR2 TOGETHER.

```

```

CC -!- SIMILARITY: Contains 8 WD repeats.
CC -!- SIMILARITY: BELONGS TO THE HIR1 FAMILY OF WD-REPEAT PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Z97204; CAB10089.1; -.
DR PIR; T40216; T40216.
DR GeneDB; SPombe; SPBC31F10.13C; -.
DR InterPro; IPR001680; WD40.
DR Pfam; PF00400; WD40; 6.
DR PRINTS; PR00320; GPROTEINRPT.
DR ProDom; PD000018; WD40; 2.
DR SMART; SM00320; WD40; 6.
DR PROSITE; PS00678; WD_REPEATS_1; 1.
DR PROSITE; PS50082; WD_REPEATS_2; 4.
DR PROSITE; PS50294; WD_REPEATS_REGION; 1.
KW Hypothetical protein; Transcription regulation; Repeat; WD repeat;
KW Nuclear protein.
FT REPEAT 16 55 WD 1.
FT REPEAT 72 111 WD 2.
FT REPEAT 132 171 WD 3.
FT REPEAT 174 213 WD 4.
FT REPEAT 268 316 WD 5.
FT REPEAT 320 361 WD 6.
FT REPEAT 623 663 WD 7.
FT REPEAT 664 709 WD 8.
SQ SEQUENCE 932 AA; 103686 MW; 09D527CDB9003D3E CRC64;

Query Match 44.2%; Score 42; DB 1; Length 932;
Best Local Similarity 46.7%; Pred. No. 40;
Matches 7; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

OY 3 SYSFIRAHSEVODLI 17
DB 126 SYRRLGSHNDIQDI 140

RESULT 13
DCLZ ECOLI STANDARD; PRT; 713 AA.
ID DCLZ ECOLI
AC P52035; P78299;
DT 01-OCT-1995 (Rel. 34, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Lysine decarboxylase, constitutive (BC 4.1.1.18) (LDC).
GN LDC OR LDC OR LDC OR B0186.
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A., PARTIAL SEQUENCE, AND CHARACTERIZATION.
RC STRAIN=K12 / W3110;
RX MEDLINE=97369816; PubMed=9226257;
RA Kirchoff Y., Kojima H., Tanaka T., Takatsuka Y., Kamio Y.;
RT "Characterization of a second lysine decarboxylase isolated from
RT Escherichia coli.";
RL J. Bacteriol. 179:4486-4492(1997).
RN [2]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 1-40 AND 415-456.
RC STRAIN=K12 / W3110;
RX MEDLINE=9748027; PubMed=9339543;
RA Yamamoto Y., Miwa Y., Miyoshi K., Furuyama J., Ohmori H.;
RT "The Escherichia coli ldc gene encodes another lysine decarboxylase,
RT probably a constitutive enzyme.";
RL Genes Genet. Syst. 72:167-172(1997).
RN [3]

```


RP SEQUENCE FROM N.A.
 RC STRAIN=K12 / MG1655;
 RX MEDLINE=97426617; PubMed=9278503;
 RA Blatner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
 RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
 RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
 RA Mau B., Shao Y.;
 RT "The complete genome sequence of *Escherichia coli* K-12.";
 RL Science 277:1453-1474(1997).
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN=K12 / MG110;
 RA Takemoto K., Mori H., Murayama N., Kataoka K., Yano M., Itoh T.,
 RA Yamamoto Y., Inokuchi H., Miki T., Hatada E., Fukuda R., Ichihara S.,
 RA Mizuno T., Makino K., Nakata A., Yura T., Samped G., Mizobuchi K.;
 RT "Systematic sequencing of the *Escherichia coli* genome: analysis of the
 RT 4.0 - 6.0 min (189,987 - 281,416bp) region.";
 RL Submitted (FEB-1996) to the EMBL/GenBank/DBJ databases.
 RN [5]
 RP SEQUENCE FROM N.A.
 RC Schramm S., Duncan M., Allen E., Araujo R., Aparicio A., Chung E.,
 RA Davis K., Federspiel N., Hyman R., Kalman S., Komp C., Kurdi O.,
 RA Lashari D., Lew H., Lin D., Namath A., Oefner P., Roberts D.,
 RA Davis R.W.;
 RL Submitted (SEP-1996) to the EMBL/GenBank/DBJ databases.
 RN [6]
 RP SEQUENCE OF 555-713 FROM N.A.
 RC STRAIN=K12;
 RX MEDLINE=98369317; PubMed=9723924;
 RA Pitchoff S., Alibaud L., Guedat A., Castanie M.-P., Bouche J.-P.;
 RT "An *Escherichia coli* gene (yaeO) suppresses temperature-sensitive
 RT mutations in essential genes by modulating Rho-dependent
 RT transcription termination.";
 RL Mol. Microbiol. 29:859-869(1998).
 CC -1- FUNCTION: IDC IS CONSTITUTIVELY BUT WEAKLY EXPRESSED UNDER VARIOUS
 CC CONDITIONS. OPTIMUM ACTIVITY IS ACHIEVED BETWEEN PH 6.2 TO 8.0.
 CC -1- CATALYTIC ACTIVITY: L-lysine = cadaverine + CO(2).
 CC -1- COFACTOR: Pyridoxal phosphate.
 CC -1- SUBUNIT: HOMODECAMER.
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF ORNITHINE, LYSINE, AND ARGININE
 CC DECARBOXYLASES.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: D87516; BAA21656.1; -;
 DR EMBL: D49445; BAA08426.1; -;
 DR EMBL: AEO00128; AAC73297.1; -;
 DR EMBL: D83536; BAA77861.1; -;
 DR EMBL: U70214; AAB08615.1; -;
 DR EMBL: Z50870; CAA90749.1; -;
 DR PIR: B64743; B64743.
 DR HSSP: P43099; IC4K.
 DR EcoGene: EG31219; lddc.
 DR InterPro: IPR000310; Decarboxylase1.
 DR InterPro: IPR005108; OKR_DC_1_N.
 DR Pfam: PF01276; OKR_DC_1_1.
 DR Pfam: PF03711; OKR_DC_1_C_1.
 DR Pfam: PF03709; OKR_DC_1_N_1.
 DR PROSITE: PS00703; OKR_DC_1_1.
 KW Lyase; Decarboxylase; Pyridoxal phosphate; Multigene family;
 KW Complete proteome.
 FT BINDING 367
 FT CONFLICT 284
 FT CONFLICT 314
 FT CONFLICT 411
 FT CONFLICT 413
 FT CONFLICT 498
 M -> T (IN REF. 2).
 N -> F (IN REF. 2).
 T -> S (IN REF. 2).
 AA -> R (IN REF. 2).
 M -> I (IN REF. 2).

FT CONFLICT 673 673 L -> I (IN REF. 2).
 SQ SEQUENCE 713 AA; 80590 MW; 4532C6069744BDF CRC64;
 Query Match 43.7%; Score 41.5; DB 1; Length 713;
 Best Local Similarity 45.0%; Pred. No. 36;
 Matches 9; Conservative 3; Mismatches 5; Indels 3; Gaps 1;
 QY 1 VPSYIFIRAH--DSVQDL 17
 DB 77 LPVAFINTHSTMDVSVDPM 96
 RESULT 14
 MIDZ BACSU STANDARD; PRT; 344 AA.
 ID MIDZ BACSU
 AC P26935;
 DT 01-AUG-1992 (Rel. 23, Created)
 DT 01-AUG-1992 (Rel. 23, Last sequence update)
 DT 26-FEB-2003 (Rel. 41, Last annotation update)
 DE Myo-inositol 2-dehydrogenase (EC 1.1.1.18).
 GN IDH OR IDLG OR E83G.
 OS Bacillus subtilis.
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 OX NCBI_TaxID=1423;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=168 / 60015;
 RX MEDLINE=92104493; PubMed=1761221;
 RA Fujita Y., Shindo K., Miwa Y., Yoshida K.;
 RT "Bacillus subtilis inositol dehydrogenase-encoding gene (idh):
 RT sequence and expression in *Escherichia coli*.";
 RL Gene 108:121-125(1991).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=168 / BGSCTA1;
 RX MEDLINE=95039891; PubMed=7952181;
 RA Yoshida K., Sano H., Miwa Y., Ogasawara N., Fujita Y.;
 RT "Cloning and nucleotide sequencing of a 15 kb region of the *Bacillus*
 RT subtilis


```

RT subtlis.";
RL Nature 390:249-256(1997).
RN [4]
RP CHARACTERIZATION.
RX MEDLINE=79239346; PubMed=112095;
RA Ramealey R., Fujita Y., Freese E.;
RT "Purification and properties of Bacillus subtilis inositol
  dehydrogenase."
RL J. Biol. Chem. 254:7684-7690(1979).
CC -1- CATALYTIC ACTIVITY: Myo-inositol + NAD(+) = 2,4,6/3,5-
  pentahydroxycyclohexanone + NADH.
CC -1- PATHWAY: Myo-inositol catabolism; first step.
CC -1- SUBUNIT: Monomer.
CC -1- INDUCTION: INOSITOL. SUBJECTED TO CATABOLITE REPRESSION.
CC -1- SIMILARITY: BELONGS TO THE GFO/IDH/MOCA FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M76431; AAA22543.1; -.
DR EMBL: D14399; BAA03296.1; -.
DR EMBL: Z99124; CAB16006.1; -.
DR PIR: JH0511; JH0511.
DR SUBLIT: BG10669; 1db.
DR InterPro: IPR000683; GFO_IDH_MOCA.
DR InterPro: IPR004104; GFO_IDH_MOCA_C.
DR Pfam: PF01408; GFO_IDH_MOCA; 1.
DR Pfam: PF02894; GFO_IDH_MOCA_C; 1.
DR Oxidoreductase; NAD; Complete proteome.
SQ SEQUENCE 344 AA; 38351 MW; 2FC908D4E2C332F CRC64;

Query Match 43.2%; Score 41; DB 1; Length 344;
Best Local Similarity 53.8%; Pred. No. 20;
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 6 FIRAHSEVODLI 18
   |||:|:|:|
Db 276 FVAAYDEIODFI 288

RESULT 15
RNB_SALTI STANDARD; PRT; 644 AA.
ID RNB_SALTI STANDARD; PRT; 644 AA.
AC Q827C9;
DT 26-FEB-2003 (Rel. 41, Created)
DT 26-FEB-2003 (Rel. 41, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Exoribonuclease II (EC 3.1.13.1) (Ribonuclease II) (RNase II).
GN RNB OR STY1350 OR T1615.
OS Salmonella typhi.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
  Enterobacteriaceae; Salmonella.
OX NCBI_TaxID=601;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=CT18;
RX MEDLINE=21534947; PubMed=11677608;
RA Parkhill J., Dougan G., James K.D., Thomson N.R., Pickard D., Wain J.,
  Churcher C., Mungall K.L., Bentley S.D., Holden M.T.G., Sebahinia M.,
  Baker S., Basham D., Brooks K., Chillingworth T., Connor P.,
  Cronin A., Davis P., Davies R.M., Dowd L., White N., Farrar J.,
  Feltham T., Hamlin N., Haque A., Hien T.T., Holroyd S., Jagers K.,
  Krogsh A., Larsen T.S., Leather S., Moule S., O'Gaora P., Parry C.,
  Quail M., Rutherford K., Simmonds M., Skelton J., Stevens K.,
  Whitehead S., Barrett B.G.;
RT "Complete genome sequence of a multiple drug resistant Salmonella
  enterica serovar Typhi CT18."
RL Nature 413:848-852(2001).

```

```

RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=TY2 / ATCC 700931;
RX MEDLINE=22531367; PubMed=12644504;
RA Deng W., Liou S.-R., Plunkett G. III, Mayhew G.F., Rose D.J.,
  Burland V., Kodyanski V., Schwartz D.C., Blattner F.R.;
RT "Comparative genomics of Salmonella enterica serovar Typhi strains Ty2
  and CT18."
RL J. Bacteriol. 185:2330-2337(2003).
CC -1- FUNCTION: Involved in mRNA degradation. Hydrolyzes single-stranded
  polyribonucleotides processively in the 3' to 5' direction (By
  similarity).
CC -1- CATALYTIC ACTIVITY: Exonucleolytic cleavage in the 3'- to 5'-
  direction to yield nucleoside 5'-phosphates.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: Belongs to the ribonuclease II (RNB) family.
CC -1- SIMILARITY: Contains 1 SI motif domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AL627270; CAD01619.1; -.
DR EMBL: AE016839; AAO69242.1; -.
DR HAMAP: MF_01036; -1.
DR InterPro: IPR004476; 3_prime_RNase.
DR InterPro: IPR002059; Cold shock.
DR InterPro: IPR001900; Ribonuclease_II.
DR InterPro: IPR003029; SI.
DR Pfam: PF00773; RNB; 1.
DR Pfam: PF00575; SI; 1.
DR SMART: SM00357; CSP; 1.
DR SMART: SM00316; SI; 1.
DR TIGRPFAM: TIGR00358; 3_prime_RNase; 1.
DR PROSITE: PS01175; RIBONUCLEASE_II; 1.
DR PROSITE: PS50126; SI; FALSE_NEG.
KW Hydrolyase; Exonuclease; Nuclease; RNA-binding; Complete proteome.
FT DOMAIN 561 643
   SI MOTIF.
SQ SEQUENCE 644 AA; 72481 MW; 17AA90ACC4C8D6B CRC64;

Query Match 43.2%; Score 41; DB 1; Length 644;
Best Local Similarity 38.9%; Pred. No. 39;
Matches 7; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

QY 2 PSYFTRAHSEVODLIA 19
   |||:|:|:|
Db 362 PDYRFLGKGEVLDIVA 379

Search completed: November 13, 2003, 09:45:27
Job time: 7.8436 secs

```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 13, 2003, 09:31:40 ; Search time 28.6351 Seconds

(without alignments)
171.224 Million cell updates/sec

Title: US-09-290-049A-2

Perfect score: 95

Sequence: 1 VPSSYFIRAHDSVEVDLIA 19

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SPTREMBL_23.*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mmc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteriophage:*
17: sp_archaea:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	84	88.4	1590	2	Q59983 streptococc
2	84	88.4	1590	2	Q55263 streptococc
3	80	84.2	1477	2	Q91466 leuconostoc
4	80	84.2	1508	2	Q95285 leuconostoc
5	80	84.2	1508	2	Q52224 leuconostoc
6	80	84.2	1527	2	Q92AR4 leuconostoc
7	80	84.2	1527	2	Q8KRE1 leuconostoc
8	80	84.2	1554	2	Q8KZL5 streptococc
9	79	83.2	1290	2	Q48756 leuconostoc
10	77	81.1	1016	2	Q91CJ7 leuconostoc
11	75	78.9	2835	2	Q8G9Q2 streptococc
12	74	77.9	1512	2	Q9WXJ5 streptococc
13	73	76.8	1338	2	Q9WXJ4 streptococc
14	66	69.5	2057	2	Q9RE05 leuconostoc
15	65	68.4	1449	2	Q68542 streptococc
16	65	68.4	1449	2	Q55264 streptococc

17	65	68.4	1518	2	Q00600 streptococc
18	64	67.4	1575	2	Q9LCH3 streptococc
19	64	67.4	1577	2	Q54178 streptococc
20	64	67.4	1599	2	Q00599 streptococc
21	61	64.2	1577	2	Q55265 streptococc
22	50	52.6	93	2	Q92IX9 streptococc
23	47	49.5	480	5	Q95WU6 bacillus lam
24	46	48.4	66	2	Q9K5K6 bacillus mo
25	46	48.4	558	5	Q9GYN4 caenorhabdi
26	46	48.4	575	5	P90900 mycoplasma
27	46	48.4	631	16	Q98P18 mycoplasma
28	45	48.4	1032	10	Q9F1J9 arabidopsis
29	46	48.4	1035	10	Q8GS71 arabidopsis
30	45.5	47.9	248	16	Q987H1 rhizobium l
31	44	46.3	249	17	Q978L3 thermoplasma
32	44	46.3	327	10	Q9C7I1 arabidopsis
33	44	46.3	371	8	Q9B8A8 trichinella
34	43	45.3	51	16	Q9RNP7 vibrio chol
35	43	45.3	64	2	Q9K5I0 bacillus su
36	43	45.3	590	17	Q9PSR0 methanobac
37	43	45.3	1260	10	Q9LMD9 arabidopsis
38	42	44.2	247	10	Q9M8T9 arabidopsis
39	42	44.2	431	2	Q47843 pantoea agg
40	42	44.2	466	12	Q8OLN2 arabidopsis
41	42	44.2	512	12	Q8JWF0 arabidopsis
42	42	44.2	563	16	Q9JZL7 neisseria m
43	42	44.2	563	16	Q9JUP8 neisseria m
44	42	44.2	644	16	Q8FHT6 escherichia
45	42	44.2	881	16	Q8G5W2 bifidobacte

ALIGNMENTS

RESULT 1	PRELIMINARY:	PRT: 1590 AA.
Q59983		
AC Q59983		
DT 01-NOV-1996 (TREMBLrel. 01, Created)		
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)		
DT 01-OCT-2002 (TREMBLrel. 22, Last annotation update)		
DE Glucosyltransferase-I precursor (EC 2.4.1.5).		
GN GTF.		
OS Streptococcus sobrinus.		
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;		
CC Streptococcus.		
CX NCBI_TaxID=1310;		
RN [1]		
RP SEQUENCE FROM N.A.		
RC STRAIN=OM2176;		
RX MEDLINE=94146405; PubMed=8312602;		
RA Sato S., Inoue M., Handa N., Aizawa Y., Isebe Y., Katayama T.;		
RT "DNA sequence of the glucosyltransferase gene of serotype d		
RT Streptococcus sobrinus."		
RL DNA Seq. 4:19-27(1993).		
DR EMBL, D13858; BAA02976.1; -		
DR InterPro; IPR002479; CW_binding.		
DR InterPro; IPR00318; Glyco_hydro_70.		
DR Pfam; PF01473; CW_binding_1; 16.		
DR Pfam; PF02324; Glyco_hydro_70; 1.		
KW Glycosyltransferase; Signal; Transferase.		
FT SIGNAL		
FT CHAIN		
SQ SEQUENCE 1590 AA; 175955 MW; C3C83A57CF3C2B0E CRC64;		
Query Match	88.4%;	Score 84; DB 2; Length 1590;
Best Local Similarity	88.9%;	Pred. No. 1.7e-05;
Matches 16; Conservative 1; Mismatches 1; Indels 0; Gaps 0;		
QY 1 VPSSYFIRAHDSVEVDLI 18		
DB 548 VPSSYFIRAHDSVEVDLI 565		

```
RESULT 2
Q55263 ID PRELIMINARY; PRT; 1590 AA.
AC O55263;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-JUN-1998 (TREMBLrel. 05, Last sequence update)
DT 01-OCT-2002 (TREMBLrel. 22, Last annotation update)
DE GTF-1.
GN GLUCOSYLTRANSFERASE.
OS Streptococcus sobrinus.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1310;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 33478;
RA Sato S.;
RT "DNA and amino-acid sequences of water-insoluble-glucan synthetase
produced from Streptococcus sobrinus ATCC 33478."
RL Ann. Kagoshima Univ. Dental School 16:23-29(1996).
DR EMBL; D63570; BAA09792.1; -.
DR InterPro; IPR002479; CW binding.
DR InterPro; IPR003318; Glyco_hydro_70.
DR Pfam; PF01473; CW_binding_1; 15.
DR Pfam; PF02324; Glyco_hydro_70; 1.
SQ SEQUENCE 1590 AA; 176058 MW; 9DF7A3F2C6EAFD43 CRC64;

Query Match 88.4%; Score 84; DB 2; Length 1590;
Best Local Similarity 88.9%; Pred. No. 1.7e-05;
Matches 16; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 VPSYSFRAHDSVQDII 18
Db 548 VPSYSFRAHDSVQDII 565

RESULT 3
Q9L466 ID PRELIMINARY; PRT; 1477 AA.
AC Q9L466;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-OCT-2002 (TREMBLrel. 22, Last annotation update)
DE Dextranucrase (EC 2.4.1.5).
GN DSRB.
OS Leuconostoc mesenteroides.
OC Bacteria; Firmicutes; Lactobacillales; Leuconostoc.
OX NCBI_TaxID=1245;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NRRL B-1355;
RA Arguello-Morales M.A., Remaud-Simeon M., Piznut S., Sarcabal P.,
Willemt R.M., Monsan P.;
RT "Sequence analysis of the gene encoding alternansucrase, a sucrose
RT glucosyltransferase from leuconostoc mesenteroides NRRL B-1355."
RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ250112; CAB76565.1; -.
DR InterPro; IPR002479; CW binding.
DR InterPro; IPR003318; Glyco_hydro_70.
DR Pfam; PF01473; CW_binding_1; 14.
DR Pfam; PF02324; Glyco_hydro_70; 1.
DR Glycosyltransferase; Transferase.
KM GLYCOSYLtransferase; Transferase.
SQ SEQUENCE 1477 AA; 164887 MW; B6F5710DEDFCB831 CRC64;

Query Match 84.2%; Score 80; DB 2; Length 1477;
Best Local Similarity 73.7%; Pred. No. 8e-05;
Matches 14; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 VPSYSFRAHDSVQDII 19
Db 603 IPNYSFVRAHDSVQTVIA 621
```

```
RESULT 4
Q9EZH5 ID PRELIMINARY; PRT; 1508 AA.
AC Q9EZH5;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-OCT-2002 (TREMBLrel. 22, Last annotation update)
DE Dextranucrase DsrB742.
GN DSRB742.
OS Leuconostoc mesenteroides.
OC Bacteria; Firmicutes; Lactobacillales; Leuconostoc.
OX NCBI_TaxID=1245;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=B-742CB;
RA Kim H.-S., Kim D., Ryu H.-J., Robyt J.F.;
RT "Leuconostoc mesenteroides B-742CB, a dextranucrase gene."
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF294469; AAG38021.1; -.
DR InterPro; IPR002479; CW binding.
DR InterPro; IPR003318; Glyco_hydro_70.
DR Pfam; PF01473; CW_binding_1; 14.
DR Pfam; PF02324; Glyco_hydro_70; 1.
DR Pfam; PF02324; Glyco_hydro_70; 1.
SQ SEQUENCE 1508 AA; 168542 MW; E2FCFA0F87AE4F3A CRC64;

Query Match 84.2%; Score 80; DB 2; Length 1508;
Best Local Similarity 73.7%; Pred. No. 8.1e-05;
Matches 14; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 VPSYSFRAHDSVQDII 19
Db 634 IPNYSFVRAHDSVQTVIA 652

RESULT 5
O52224 ID PRELIMINARY; PRT; 1508 AA.
AC O52224;
DT 01-JUN-1998 (TREMBLrel. 06, Created)
DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)
DT 01-OCT-2002 (TREMBLrel. 22, Last annotation update)
DE Glucosyltransferase (EC 2.4.1.5).
GN DSRB.
OS Leuconostoc mesenteroides.
OC Bacteria; Firmicutes; Lactobacillales; Leuconostoc.
OX NCBI_TaxID=1245;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NRRL B-1299;
RA Monchois V., Remaud-Simeon M., Monsan P., Willemt R.M.;
RT "Cloning and sequencing of a gene coding for an extracellular
RT dextranucrase (DSRB) from Leuconostoc mesenteroides NRRL B-1299
RT synthesizing only a a(1-6) glucan."
RL FEMS Microbiol. Lett. 0:0-0(1998).
DR EMBL; AF030129; AAB95453.1; -.
DR InterPro; IPR002479; CW binding.
DR InterPro; IPR003318; Glyco_hydro_70.
DR Pfam; PF01473; CW_binding_1; 14.
DR Pfam; PF02324; Glyco_hydro_70; 1.
DR Glycosyltransferase; Transferase.
KM GLYCOSYLtransferase; Transferase.
SQ SEQUENCE 1508 AA; 168511 MW; E70CECB57A70D1F0 CRC64;

Query Match 84.2%; Score 80; DB 2; Length 1508;
Best Local Similarity 73.7%; Pred. No. 8.1e-05;
Matches 14; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 VPSYSFRAHDSVQDII 19
Db 634 IPNYSFVRAHDSVQTVIA 652

RESULT 6
```

```

Q9ZAR4          PRELIMINARY;      PRT; 1527 AA.
ID  Q9ZAR4;
AC  Q9ZAR4;
DT  01-MAY-1999 (TREMBLrel. 10, Created)
DE  01-MAY-1999 (TREMBLrel. 10, Last sequence update)
DT  01-OCT-2002 (TREMBLrel. 22, Last annotation update)
DE  Dextranucrase.
GN  DEX.
OS  Leuconostoc mesenteroides.
OC  Bacteria; Firmicutes; Lactobacillales; Leuconostoc.
OX  NCBI_TaxID=1245;
RN  [1]
RP  SEQUENCE FROM N.A.
RC  STRAIN=NRRL B-512-F;
RA  Bhattacharya R., Singh D.K.S.;
RT  "Cloning and Molecular Characterization of Dextranucrase Gene from
RT  Leuconostoc mesenteroides NRRL B-512F."
RT  Submitted (DEC-1996) to the EMBL/GenBank/DBJ databases.
DR  EMBL; 081374; AAD10952.1; -
DR  InterPro; IPR002479; CW_binding.
DR  InterPro; IPR003318; Glyco_hydro_70.
DR  Pfam; PF01473; CW_binding_1; 16.
DR  Pfam; PF02324; Glyco_hydro_70; 1.
SQ  SEQUENCE 1527 AA; 169709 MW; 1DFAFA237C743398 CRC64;

Query Match          84.2%; Score 80; DB 2; Length 1527;
Best Local Similarity 73.7%; Pred. No. 8.3e-05;
Matches 14; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY  1 VPSYSFIRAHDSVQDLIA 19
Db  652 IPNYSFVRHDSVQTVIA 670

RESULT 7
Q8KRE1          PRELIMINARY;      PRT; 1527 AA.
ID  Q8KRE1;
AC  Q8KRE1;
DT  01-OCT-2002 (TREMBLrel. 22, Created)
DT  01-OCT-2002 (TREMBLrel. 22, Last sequence update)
DT  01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE  Dextranucrase Dsrd (EC 2.4.1.5).
GN  DSRD.
OS  Leuconostoc mesenteroides.
OC  Bacteria; Firmicutes; Lactobacillales; Leuconostoc.
OX  NCBI_TaxID=1245;
RN  [1]
RP  SEQUENCE FROM N.A.
RA  Neubauer H., Bauche A., Mollet B.;
RT  "Isolation and characterization of the dextranucrase Dsrd of
RT  Leuconostoc mesenteroides Lcc4."
RT  Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
DR  EMBL; AY017384; AAG61158.1; -
DR  InterPro; IPR002479; CW_binding.
DR  InterPro; IPR003318; Glyco_hydro_70.
DR  Pfam; PF01473; CW_binding_1; 6.
DR  Pfam; PF02324; Glyco_hydro_70; 1.
DR  Pfam; PF02324; Glycosyltransferase.
KM  TRANSFERASE; Glycosyltransferase.
SQ  SEQUENCE 1527 AA; 169835 MW; F9D0DE220BD89668 CRC64;

Query Match          84.2%; Score 80; DB 2; Length 1527;
Best Local Similarity 73.7%; Pred. No. 8.3e-05;
Matches 14; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY  1 VPSYSFIRAHDSVQDLIA 19
Db  652 IPNYSFVRHDSVQTVIA 670

RESULT 8
Q8KZL5          PRELIMINARY;      PRT; 1554 AA.
ID  Q8KZL5;
AC  Q8KZL5;

```

```

DT  01-OCT-2002 (TREMBLrel. 22, Created)
DT  01-OCT-2002 (TREMBLrel. 22, Last sequence update)
DT  01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE  Glucosyltransferase.
GN  GTFU.
OS  Streptococcus sobrinus.
OC  Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC  Streptococcus.
OX  NCBI_TaxID=1310;
RN  [1]
RP  SEQUENCE FROM N.A.
RX  MEDLINE=21956884; PubMed=11960691;
RA  Hanada N., Fukushima K., Nomura Y., Sengoku H., Hayakawa M.,
RA  Mukasa H., Shirota T., Abiko Y.;
RT  "Cloning and nucleotide sequence analysis of the Streptococcus
RT  sobrinus gltF gene that produces a highly branched water-soluble
RT  glucan."
RT  Biochim. Biophys. Acta 1570:75-79(2002).
DR  EMBL; AB089438; BAC07265.1; -
DR  InterPro; IPR002479; CW_binding.
DR  InterPro; IPR003318; Glyco_hydro_70.
DR  Pfam; PF01473; CW_binding_1; 6.
DR  Pfam; PF02324; Glyco_hydro_70; 1.
KM  TRANSFERASE.
SQ  SEQUENCE 1554 AA; 171676 MW; 6981BCC1DAE24A73 CRC64;

Query Match          84.2%; Score 80; DB 2; Length 1554;
Best Local Similarity 73.7%; Pred. No. 8.4e-05;
Matches 14; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY  1 VPSYSFIRAHDSVQDLIA 19
Db  557 IPTISFVRHDSVQTVIA 575

RESULT 9
Q48756          PRELIMINARY;      PRT; 1290 AA.
ID  Q48756;
AC  Q48756;
DT  01-NOV-1996 (TREMBLrel. 01, Created)
DT  01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT  01-OCT-2002 (TREMBLrel. 22, Last annotation update)
DE  Dextranucrase.
OS  Leuconostoc mesenteroides.
OC  Bacteria; Firmicutes; Lactobacillales; Leuconostoc.
OX  NCBI_TaxID=1245;
RN  [1]
RP  SEQUENCE FROM N.A.
RC  STRAIN=NRRL B1299;
RX  MEDLINE=97136686; PubMed=8982063;
RA  Monchois V., Willemot R.M., Renaud-Simeon M., Croux C., Monsan P.;
RT  "Cloning and sequencing of a gene coding for a novel dextranucrase
RT  from Leuconostoc mesenteroides NRRL B-1299 synthesizing only alpha (1-
RT  6) and alpha (1-3) linkages."
RL  Gene 182:23-32(1996).
DR  EMBL; U38181; AAB40875.1; -
DR  InterPro; IPR002479; CW_binding.
DR  InterPro; IPR003318; Glyco_hydro_70.
DR  Pfam; PF01473; CW_binding_1; 11.
DR  Pfam; PF02324; Glyco_hydro_70; 1.
SQ  SEQUENCE 1290 AA; 145590 MW; 3555C2E96B749FAA CRC64;

Query Match          83.2%; Score 79; DB 2; Length 1290;
Best Local Similarity 83.3%; Pred. No. 0.0001;
Matches 15; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY  2 PYSYFIRAHDSVQDLIA 19
Db  388 PNYSFVRHDSVQTVIA 405

RESULT 10
Q9LCJ7

```

ID Q9LCU7 PRELIMINARY; PRT; 1016 AA.
AC Q9LCU7;
DE 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DE 01-OCT-2002 (Tremblrel. 22, Last annotation update)
DE Dextranucrase.
GN DSRT.
OS Leuconostoc mesenteroides.
OC Bacteria; Firmicutes; Lactobacillales; Leuconostoc.
NCBI_TaxID=1245;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=NRRL B-512P;
RX MEDLINE=20169623; PubMed=10705445; Kobayashi M.;
Funane K., Mizuno K., Takahara H.;
RT "Gene encoding a dextranucrase-like protein in Leuconostoc
mesenteroides NRRL B-512P."
RL Bioecl. Biotechnol. Biochem. 64:29-38(2000).
DR EMBL; AB020020; BAA90527.1; -.
DR HSSP; P06278; IVIS.
DR InterPro; IPR003318; Glyco_hydro_70.
DR Pfam; PF02324; Glyco_hydro_70; 1.
SQ SEQUENCE 1016 AA; 110344 MW; 8896EFDE13CCCB47 CRC64;

Query Match 81.1%; Score 77; DB 2; Length 1016;
Best Local Similarity 77.8%; Pred. No. 0.00018;
Matches 14; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
QY 2 PSYSFIRAHDSSEVQDIIA 19
DB 625 PNYSFVRAHDSSEVQTVIA 642

RESULT 11
ID Q8G9Q2 PRELIMINARY; PRT; 2835 AA.
AC Q8G9Q2;
DE 01-MAR-2003 (Tremblrel. 23, Created)
DT 01-MAR-2003 (Tremblrel. 23, Last sequence update)
DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
DE Dextranucrase (EC 2.4.1.5) (Fragment).
GN DSRE.
OS Leuconostoc mesenteroides.
OC Bacteria; Firmicutes; Lactobacillales; Leuconostoc.
NCBI_TaxID=1245;
RN (1)
RP SEQUENCE FROM N.A.
RX MEDLINE=2231661; PubMed=12270834;
RA Bozomet S., Dols-Lafargue M., Fabre E., Pizzut S., Remaud-Simeon M.,
Mondan P., Willemot R.M.;
RT "Molecular characterisation of DSR-E, an alpha-1,2 linkage
synthesising dextranucrase with two catalytic domains."
RL U. Bacteriol. 184:5753-5761(2002).
DR EMBL; A0430204; CAD2883.1; -.
DR Transferase; Glycosyltransferase.
FT NON TER 1
SQ SEQUENCE 2835 AA; 313264 MW; D03262CDD735399D CRC64;

Query Match 78.9%; Score 75; DB 2; Length 2835;
Best Local Similarity 77.8%; Pred. No. 0.0012;
Matches 14; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 2 PSYSFIRAHDSSEVQDIIA 19
DB 623 PNYSFVRAHDSSEVQTVIA 646

RESULT 12
Q9WXJ5 PRELIMINARY; PRT; 1512 AA.
ID Q9WXJ5
AC Q9WXJ5;
DE 01-NOV-1999 (Tremblrel. 12, Created)
DT 01-NOV-1999 (Tremblrel. 12, Last sequence update)

DT 01-OCT-2002 (Tremblrel. 22, Last annotation update)
DE GTF-S.
GN GTF.
OS Streptococcus criceti.
OC Plasmid pAM1.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
NCBI_TaxID=1333;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=HS-6;
RA Inoue M., Fukui K., Miyagi A.;
RT "S.cricetus glucosyltransferase (gifs and gtf) genes."
RL Submitted (MAR-1999) to the EMBL/Genbank/DBJ databases.
DR EMBL; AB026123; BAA77237.1; -.
DR InterPro; IPR002479; CW_binding.
DR InterPro; IPR003318; Glyco_hydro_70.
DR Pfam; PF01473; CW_binding_1; 14.
DR Pfam; PF02324; Glyco_hydro_70; 1.
KW Plasmid.
SQ SEQUENCE 1512 AA; 167145 MW; 4C03D9C8C601FC14 CRC64;

Query Match 77.9%; Score 74; DB 2; Length 1512;
Best Local Similarity 77.8%; Pred. No. 0.00091;
Matches 14; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
QY 2 PSYSFIRAHDSSEVQDIIA 19
DB 560 PSYVFRVRAHDSSEVQTVIA 577

RESULT 13
ID Q9WXJ4 PRELIMINARY; PRT; 1338 AA.
AC Q9WXJ4;
DE 01-NOV-1999 (Tremblrel. 12, Created)
DT 01-NOV-1999 (Tremblrel. 12, Last sequence update)
DT 01-OCT-2002 (Tremblrel. 22, Last annotation update)
DE GTF-S.
GN GTF.
OS Streptococcus criceti.
OC Plasmid pAM1.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
NCBI_TaxID=1333;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=HS-6;
RA Inoue M., Fukui K., Miyagi A.;
RT "S.cricetus glucosyltransferase (gifs and gtf) genes."
RL Submitted (MAR-1999) to the EMBL/Genbank/DBJ databases.
DR EMBL; AB026123; BAA77236.1; -.
DR InterPro; IPR002479; CW_binding.
DR InterPro; IPR003318; Glyco_hydro_70.
DR Pfam; PF01473; CW_binding_1; 10.
DR Pfam; PF02324; Glyco_hydro_70; 1.
KW Plasmid.
SQ SEQUENCE 1338 AA; 148558 MW; 0A90C8E10E15D99B CRC64;

Query Match 76.8%; Score 73; DB 2; Length 1338;
Best Local Similarity 78.9%; Pred. No. 0.0012;
Matches 15; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 VPSYFIRAHDSSEVQDIIA 19
DB 509 VPSYFIRAHDSSEVQTVIA 527

RESULT 14
Q9RE05 PRELIMINARY; PRT; 2057 AA.
ID Q9RE05
AC Q9RE05;
DE 01-MAY-2000 (Tremblrel. 13, Created)

DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-OCT-2002 (TREMBLrel. 22, Last annotation update)
DE Alternansucrase (EC 2.4.1.140).
GN ASR.
OS Leuconostoc mesenteroides.
OC Bacteria; Firmicutes; Lactobacillales; Leuconostoc.
OX NCBI_TaxID=1245;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NRRL B-1355.
RX MEDLINE=20080809; PubMed=10612736;
RA Arcuello-Morales M.A., Renaud-Simeon M., Pizatz S., Sarrabal P.,
RA Willemot R.M., Monsan P.;
RT "Sequence analysis of the gene encoding alternansucrase, a sucrose
RT glucosyltransferase from leuconostoc mesenteroides NRRL B-1355.";
RL FEMS Microbiol. Lett. 182:81-85(2000).
DB EMBL; AJ250173; CAB65910.2; -;
DR InterPro; IPR002479; CW_binding.
DR InterPro; IPR003318; Glyco_Hydro_70.
DR Pfam; PF01473; CW_binding_1; 12.
DR Pfam; PF02324; Glyco_Hydro_70; 1.
KW Glycosyltransferase; Transferase.
SQ SEQUENCE 2057 AA; 228987 MW; 62BC93385D9A11BE CRC64;

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 13, 2003, 09:31:40 ; Search time 40.3507 Seconds
(without alignments)
86.541 Million cell updates/sec

Title: us-09-290-049a-3

Sequence: 1 TGAATGQLLFRANGVQVKG 22

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

```
1: A_Geneseq_19Jun03:*
2: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
3: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
4: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
5: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
6: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
7: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
8: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
9: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
10: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
11: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
12: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
13: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
14: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
15: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
16: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
17: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
18: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
19: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
20: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
21: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
22: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
23: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*
24: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*
25: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2003.DAT:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	112	100.0	1475	23	AAU98027
2	112	100.0	1475	23	AAU98030
3	112	100.0	1475	23	AAU98031
4	112	100.0	1475	23	AAU98032
5	112	100.0	1475	23	AAU98033
6	112	100.0	1475	23	AAU98034
7	112	100.0	1475	23	AAU98035
8	112	100.0	1475	23	AAU98036
9	112	100.0	1475	23	AAU98037

10	112	100.0	1475	23	AAU98038
11	112	100.0	1475	23	AAU98039
12	112	100.0	1475	23	AAU98040
13	105	93.8	380	23	AAU79286
14	105	93.8	1476	23	AAU79284
15	87	77.7	22	14	AAU43696
16	87	77.7	22	18	AAU44159
17	87	77.7	22	18	AAU43358
18	87	77.7	24	18	AAU43164
19	87	77.7	236	14	ABP56632
20	87	77.7	1592	14	ABP32925
21	81	72.3	1835	23	ABP98574
22	78	69.6	1375	23	AAU98028
23	78	69.6	1375	23	AAU79288
24	67	59.8	1430	23	AAU98029
25	67	59.8	1430	23	AAU98041
26	67	59.8	1430	23	AAU98042
27	67	59.8	1430	23	AAU98043
28	67	59.8	1430	23	AAU98044
29	67	59.8	1430	23	AAU98045
30	66	58.9	1577	17	AAU91047
31	65	58.0	2057	21	AAU10667
32	57	50.9	1527	23	AAU80055
33	54.5	48.7	32	23	ABP98593
34	49.5	44.2	320	20	AAU87969
35	49.5	44.2	457	20	AAU87970
36	49.5	44.2	811	17	AAU95014
37	49.5	44.2	812	17	AAU95017
38	49.5	44.2	862	20	AAU87971
39	49.5	44.2	866	21	AAU830543
40	49.5	44.2	866	21	AAU80546
41	49.5	44.2	2710	19	AAU95016
42	49.5	44.2	2710	19	AAU68387
43	49	43.8	337	22	AAU61855
44	49	43.8	343	22	AAU1855
45	47	42.0	785	22	AAU39393

ALIGNMENTS

```
RESULT 1
AAU98027
ID AAU98027 standard; Protein; 1475 AA.
XX
AC AAU98027;
XX
DT 27-AUG-2002 (first entry)
XX
DE S. mutans glucosyltransferase GTFB.
XX
KW Glucosyltransferase; GTFB; transgenic plant; paper sizing;
KW coating composition; glucan; starch; latex; thermoplastic molecule;
KW amyloplast; vacuole; paper manufacture.
XX
OS Streptococcus mutans.
XX
PN US2002031826-A1.
XX
PD 14-MAR-2002.
XX
PF 19-DEC-2000; 2000US-0740274.
XX
PR 11-DEC-1998; 98US-0210361.
PR 07-JUN-1995; 95US-0478704.
PR 07-JUN-1995; 95US-0482711.
PR 07-JUN-1995; 95US-0485243.
PR 16-JAN-1998; 98US-0007999.
PR 16-JAN-1998; 98US-0008172.
PR 20-JAN-1998; 98US-0009620.
XX
PA (NICH/) NICHOLS S E.
XX
```

S. mutans GTFB mut
S. mutans glucosyl
S. mutans glucosyl
Streptococcus muta
Streptococcus muta
GTF subsequence. S
GTF antigenic pep
GTF antigenic epit
GTF antigenic pep
Carbohydrate strep
Glucosyltransferase
Dextran saccharase
S. mutans glucosyl
Streptococcus muta
S. mutans glucosyl
S. mutans glucosyl
S. mutans glucosyl
S. mutans glucosyl
S. mutans glucosyl
S. mutans glucosyl
Alpha-D-glucosyltr
L. mesenteroides a
Leuconostoc mesent
Dextran-saccharas
Toxin A immunogeni
Toxin A immunogeni
C. difficile toxin
C. difficile toxin
Toxin A immunogeni
A recombinant prot
C. difficile toxin
Clostridium difficile
P. falciparum yf8b
Human polypeptide

Pt	Nichols SE;
DR	WPI; 2002-414332/44.
XX	N-PSDB; ABR52938.
XX	
Pt	Glucosyltransferase B or D protein useful for producing a glucan useful
Pt	as substitutes for and additions to modified starch and latexes in
Pt	paper manufacture, comprises mutations in specific positions -
PS	Disclosure; Page 21-25; 44pp; English.
XX	
CC	The invention an isolated protein comprising a glucosyltransferase
CC	(GTF) B polypeptide having changes at position from I448V, D457N,
CC	D567T, K1014T, D457N/D567T, D457N/D571K, D567T/D571K,
CC	D567T/D571K/K1014T, I448V/D457N/D567T/D571K/K779Q/K1014T,
CC	I159A/Y170A/V171A, and K779Q or a GTF D polypeptide having
CC	changes at positions from T589D, T589E, N471D, N471D/T589D, and
CC	N471D/T589E. Also included are a glucan produced by the GTF mutant,
CC	an isolated polynucleotide which encodes P1 or P2, or its complementary
CC	polynucleotide, a ribonucleic acid sequence encoding the GTF mutant,
CC	an expression cassette comprising the polynucleotide operably linked to a
CC	promoter, a vector comprising the expression cassette, host cell
CC	introduced with the vector, a transgenic plant comprising the
CC	vector, a seed or tuber from the transgenic plant, a paper sizing and/or
CC	coating composition comprising a glucan produced in a plant transformed
CC	with a gene encoding the mutant GTF, wild type or, starch, a latex,
CC	thermoplastic molecule or their combinations or glucan and starch where
CC	the glucan is produced in the amyloplast and/or vacuole or a maize line
CC	deficient in starch biosynthesis, transformed with a gene encoding a
CC	glucosyltransferase B or D enzyme, wild-type or mutant and a paper
CC	comprising the glucan (paper sizing/coating agent). The vector is useful
CC	for producing a glucan in a plant. The method comprises transforming a
CC	plant cell with the vector, growing the plant cell under plant growing
CC	conditions to produce a regenerated plant and inducing expression of the
CC	polynucleotide for a time sufficient to produce the glucan in the
CC	regenerated plant, where the vector contains a transit sequence from
CC	ribulose biphosphate carboxylase small subunit, waxy, brittle-1 and
CC	chlorophyll AB binding protein to produce a transgenic plant, and glucan
CC	is produced in the amyloplast of potato or the vacuole of sugar beet.
CC	Glucans are useful as substitutes for and additions to modified starch
CC	and latexes in paper manufacture. Unlike prior art techniques, which
CC	require input materials that produce chemical effluents, paper
CC	manufacture utilising the glucan produced by GTF, which utilises
CC	biologically produced input materials, is more cost-effective and
CC	environmentally friendly. Moreover, glucans also exhibit thermoplastic
CC	properties and impart gloss to the paper during coating step.
CC	The present sequence represents GTFB.
XX	
SQ	Sequence 1475 AA;
OY	1 TGARTINGQLTYFRANGVOVKG 22
Db	1300 TGARTINGQLTYFRANGVOVKG 1321
	Query Match 100.0%; Score 112; DB 23; Length 1475;
	Best Local Similarity 100.0%; Pred. No. 3.8e-10;
	Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
ID	AU098030 standard; Protein; 1475 AA.
XX	AU098030;
XX	27-AUG--2002 (first entry)
XX	
DE	5. mutans glucosyltransferase GTFB mutant I448V.
KM	Glucosyltransferase; GTFB; transgenic plant; paper sizing;
KM	coating composition; glucan; starch; latex; thermoplastic molecule;
KM	amyloplast; vacuole; paper manufacture; mutant; mteuin.
XX	

OS	Streptococcus mutans.
XX	Synthetic.
XX	
FT	Key Location/Qualifiers
XX	Mis-difference 448
XX	/note= "Wild-type Ile substituted by Val"
XX	
PM	US2002031826-A1.
XX	
PD	14-MAR-2002.
XX	
PF	19-DEC-2000; 2000US-0740274.
XX	
PR	11-DEC-1998; 98US-0210361.
PR	07-JUN-1995; 95US-0478704.
PR	07-JUN-1995; 95US-0482711.
PR	07-JUN-1995; 95US-0485243.
PR	16-JAN-1998; 96US-0007999.
PR	16-JAN-1998; 96US-0008172.
PR	20-JAN-1998; 98US-0009620.
XX	
PA	(NICH) / NICHOLS S E.
XX	
EI	Nichols SE;
DR	WPI, 2002-414332/44.
XX	
PT	Glucosyltransferase B or D protein useful for producing a glucan useful as substitutes for and additions to modified starch and latexes in paper manufacture, comprises mutations in specific positions - Claim 36; Page -: 44pp; English.
PS	
XX	The invention comprising an isolated protein comprising a glucosyltransferase (GTF). B polypeptide having changes at position from I448V, D457N, D567T, K1014T, D457N/D567T, D457N/D571K, D567T/D571K, D567T/D571K/K1014T, I448V/D457N/D567T/D571K/K779Q/K1014T, Y169A/Y170Q/Y171A, and K779Q or a GTF D polypeptide having changes at positions from T589D, Y589E, N471D, N471D/T589D, and N471D/Y589E. Also included are a glucan produced by the GTF mutant, an isolated polynucleotide which encodes P1 or P2, or its complementary polynucleotide, a ribonucleic acid sequence encoding the GTF mutant, an expression cassette comprising the polynucleotide operably linked to a promoter, a vector comprising the expression cassette, host cell introduced with the vector, a transgenic plant comprising the vector, a seed or tuber from the transgenic plant, a paper sizing and/or coating composition comprising a glucan produced in a plant transformed with a gene encoding the mutant GTF, wild type or, starch, a latex, thermoelastic molecule or their combinations or glucan and starch where the glucan is produced in the amyloplast and/or vacuole or a maize line deficient in starch biosynthesis, transformed with a gene encoding a glucosyltransferase B or D enzyme, wild-type or mutant and a paper comprising the glucan (paper sizing/coating agent). The vector is useful for producing a glucan in a plant. The method comprises transforming a plant cell with the vector, growing the plant cell under plant growing conditions to produce a regenerated plant and inducing expression of the polynucleotide for a time sufficient to produce the glucan in the regenerated plant, where the vector contains a transit sequence from ribulose biphosphate carboxylase small subunit, waxy, brittle-1 and chlorophyll AB binding protein to produce a transgenic plant, and glucan is produced in the amyloplast of potato or the vacuole of sugar beet. CC Glucans are useful as substitutes for and additions to modified starch and latexes in paper manufacture. Unlike prior art techniques, which require input materials that produce chemical effluents, paper manufacture utilizing the glucan produced by GTF, which utilizes biologically produced input materials, is more cost-effective and environmentally friendly. Moreover, glucans also exhibit thermoelastic properties and impart gloss to the paper during coating step. CC The present sequence represents a GTFB mutant of the invention. CC Note: The present sequence is not shown in the specification but was created by the indexer using the GTFB sequence appearing as AU98027 and the information in Claim 36.

SQ Sequence 1475 AA;

Query Match 100.0%; Score 112; DB 23; Length 1475;
 Best Local Similarity 100.0%; Pred. No. 3.8e-10;
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TGARTINGQLLYFRANGVQVKG 22
 |||||
 1300 TGARTINGQLLYFRANGVQVKG 1321

Db

RESULT 3
 AAU98031
 ID AAU98031 standard; Protein; 1475 AA.

XX AC AAU98031;
 XX DT 27-AUG-2002 (first entry)
 XX DE S. mutans glucosyltransferase GTFB mutant D457N.
 XX KM Glucosyltransferase; GTFB; transgenic plant; paper sizing;
 XX KM coating composition; glucan; starch; latex; thermoplastic molecule;
 XX KW amyloplast; vacuole; paper manufacture; mutant; mutain.
 XX OS Streptococcus mutans.
 XX OS Synthetic.
 XX XX
 XX FH Key Location/Qualifiers
 XX FT Misc-difference 457 /note= "Wild-type Asp substituted by Asn"
 XX XX
 XX PN US2002031826-A1.
 XX XX
 XX PD 14-MAR-2002.
 XX PF 19-DEC-2000; 2000US-0740274.
 XX XX
 XX PR 11-DEC-1998; 98US-0210361.
 XX PR 07-JUN-1995; 95US-0478704.
 XX PR 07-JUN-1995; 95US-0482711.
 XX PR 07-JUN-1995; 95US-0485243.
 XX PR 16-JAN-1998; 98US-0007999.
 XX PR 16-JAN-1998; 98US-0008172.
 XX PR 20-JAN-1998; 98US-0009620.
 XX XX
 XX PA (NICH/) NICHOLS S E.
 XX XX
 XX PI Nichols SE;
 XX XX
 XX DR WPI; 2002-414332/44.
 XX XX
 XX PT Glucosyltransferase B or D protein useful for producing a glucan useful
 XX PT as substitutes for and additions to modified starch and latexes in
 XX PT paper manufacture, comprises mutations in specific positions -
 XX XX
 XX PS Claim 36; Page -; 44pp; English.
 XX XX
 XX CC The invention an isolated protein comprising a glucosyltransferase
 XX CC (GTF) B polypeptide having changes at position from 1448V, D457N,
 XX CC D567T, K1014T, D457N/D567T, D457N/D571K, D567T/D571K,
 XX CC D567T/D571K/K1014T, 1448V/D457N/D571K/K790C/K1014T,
 XX CC Y169A/Y170A/Y171A, and K779Q or a GTF D polypeptide having
 XX CC changes at positions from T589P, T589E, N471D, N471D/T589D, and
 XX CC N471D/T589E. Also included are a glucan produced by the GTF mutant,
 XX CC an isolated polynucleotide which encodes P1 or P2, or its complementary
 XX CC polynucleotide, a ribonucleic acid sequence encoding the GTF mutant,
 XX CC an expression cassette comprising the polynucleotide operably linked to a
 XX CC promoter, a vector comprising the expression cassette, host cell
 XX CC introduced with the vector, a transgenic plant comprising the
 XX CC vector, a seed or tuber from the transgenic plant, a paper sizing and/or
 XX CC coating composition comprising a glucan produced in a plant transformed
 XX CC with a gene encoding the mutant GTF, wild type or, starch, a latex,

CC thermoplastic molecule or their combinations or glucan and starch where
 CC the glucan is produced in the amyloplast and/or vacuole or a maize line
 CC deficient in starch biosynthesis, transformed with a gene encoding a
 CC glucosyltransferase B or D enzyme, wild-type or mutant and a paper
 CC comprising the glucan (paper sizing/coating agent). The vector is useful
 CC for producing a glucan in a plant. The method comprises transforming a
 CC plant cell with the vector, growing the plant cell under plant growing
 CC conditions to produce a regenerated plant and inducing expression of the
 CC polynucleotide for a time sufficient to produce the glucan in the
 CC regenerated plant, where the vector contains a transit sequence from
 CC ribulose biphosphate carboxylase small subunit, waxy, brittle-1 and
 CC chlorophyll AB binding protein to produce a transgenic plant, and glucan
 CC is produced in the amyloplast of potato or the vacuole of sugar beet.
 CC Glucans are useful as substitutes for and additions to modified starch
 CC and latexes in paper manufacture. Unlike prior art techniques, which
 CC require input materials that produce chemical effluents, paper
 CC manufacture utilizing the glucan produced by GTF, which utilizes
 CC biologically produced input materials, is more cost-effective and
 CC environmentally friendly. Moreover, glucans also exhibit thermoplastic
 CC properties and impart gloss to the paper during coating step.
 CC The present sequence represents a GTFB mutant of the invention.
 CC Note: The present sequence is not shown in the specification but
 CC was created by the indexer using the GTFB sequence appearing as AAU98027
 CC and the information in claim 36.

SQ Sequence 1475 AA;

Query Match 100.0%; Score 112; DB 23; Length 1475;
 Best Local Similarity 100.0%; Pred. No. 3.8e-10;
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TGARTINGQLLYFRANGVQVKG 22
 |||||
 1300 TGARTINGQLLYFRANGVQVKG 1321

Db

RESULT 4
 AAU98032
 ID AAU98032 standard; Protein; 1475 AA.

XX AC AAU98032;
 XX DT 27-AUG-2002 (first entry)
 XX DE S. mutans glucosyltransferase GTFB mutant D567T.
 XX KM Glucosyltransferase; GTFB; transgenic plant; paper sizing;
 XX KM coating composition; glucan; starch; latex; thermoplastic molecule;
 XX KW amyloplast; vacuole; paper manufacture; mutant; mutain.
 XX OS Streptococcus mutans.
 XX OS Synthetic.
 XX XX
 XX FH Key Location/Qualifiers
 XX FT Misc-difference 567 /note= "Wild-type Asp substituted by Thr"
 XX XX
 XX PN US2002031826-A1.
 XX XX
 XX PD 14-MAR-2002.
 XX PF 19-DEC-2000; 2000US-0740274.
 XX XX
 XX PR 11-DEC-1998; 98US-0210361.
 XX PR 07-JUN-1995; 95US-0478704.
 XX PR 07-JUN-1995; 95US-0482711.
 XX PR 07-JUN-1995; 95US-0485243.
 XX PR 16-JAN-1998; 98US-0007999.
 XX PR 16-JAN-1998; 98US-0008172.
 XX PR 20-JAN-1998; 98US-0009620.
 XX XX
 XX PA (NICH/) NICHOLS S E.
 XX XX

CC and the information in claim 36.
 XX Sequence 1475 AA;
 SQ
 Query Match 100.0%; Score 112; DB 23; Length 1475;
 Best Local Similarity 100.0%; Pred. No. 3.8e-10;
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 TGARTINGQLLYPRANGVQVKG 22
 Db 1300 TGARTINGQLLYPRANGVQVKG 1321
 RESULT 6
 AAU98034
 ID AAU98034 standard; Protein: 1475 AA.
 AC
 AAU98034;
 XX
 27-AUG-2002 (first entry)
 DT
 XX
 S. mutans glucosyltransferase GTFB mutant D457N/D567T.
 DB
 XX
 Glucosyltransferase; GTFB; transgenic plant; paper sizing;
 KM coating composition; glucan; starch; latex; thermoplastic molecule;
 KW amyloplast; vacuole; paper manufacture; mutant; mutain.
 XX
 OS Streptococcus mutans.
 OS Synthetic.
 XX
 Key Location/Qualifiers
 FH Misc-difference 457
 FT /note= "Wild-type Asp substituted by Asn"
 FT Misc-difference 567 /note= "Wild-type Asp substituted by Thr"
 FT
 FT
 XX US2002031826-A1.
 XX
 PD 14-MAR-2002.
 XX
 19-DEC-2000; 2000US-0740274.
 XX
 11-DEC-1998; 98US-0210361.
 PR 07-JUN-1995; 95US-0478704.
 PR 07-JUN-1995; 95US-0482711.
 PR 07-JUN-1995; 95US-0485243.
 PR 16-JAN-1998; 98US-0007999.
 PR 16-JAN-1998; 98US-0008172.
 PR 20-JAN-1998; 98US-0009620.
 XX
 PA (NICH/) NICHOLS S E.
 FI Nichols SE;
 XX
 DR WPI; 2002-414332/44.
 XX
 PT Glucosyltransferase B or D protein useful for producing a glucan useful
 PT as substitutes for and additions to modified starch and latexes in
 PT paper manufacture, comprises mutations in specific positions -
 XX
 PS Claim 36; Page -; 44pp; English.
 XX
 CC The invention an isolated protein comprising a glucosyltransferase
 CC (GTF) B polypeptide having changes at position from 1448V, D457N,
 CC D567T, K1014T, D457N/D567T, D457N/D571K, D567T/D571K,
 CC D567T/D571K/K1014T, 1448V/D457N/D567T/D571K/K779Q/K1014T,
 CC Y169A/Y170A/Y171A, and K779Q or a GTF D polypeptide having
 CC changes at positions from T589D, T589E, N471D, N471D/T589D, and
 CC N471D/T589E. Also included are a glucan produced by the GTF mutant,
 CC an isolated polynucleotide which encodes P1 or P2, or its complementary
 CC polynucleotide, a ribonucleic acid sequence encoding the GTF mutant,
 CC an expression cassette comprising the polynucleotide operably linked to a
 CC promoter, a vector comprising the expression cassette, host cell

CC introduced with the vector, a transgenic plant comprising the
 CC vector, a seed or tuber from the transgenic plant, a paper sizing and/or
 CC coating composition comprising a glucan produced in a plant transformed
 CC with a gene encoding the mutant GTF, wild type or, starch, a latex.
 CC thermoplastic molecule or their combinations or glucan and starch where
 CC the glucan is produced in the amyloplast and/or vacuole or a maize line
 CC deficient in starch biosynthesis, transformed with a gene encoding a
 CC glucosyltransferase B or D enzyme, wild-type or mutant and a paper
 CC comprising the glucan (paper sizing/coating agent). The vector is useful
 CC for producing a glucan in a plant. The method comprises transforming a
 CC plant cell with the vector, growing the plant cell under plant growing
 CC conditions to produce a regenerated plant and inducing expression of the
 CC polynucleotide for a time sufficient to produce the glucan in the
 CC regenerated plant, where the vector contains a transit sequence from
 CC chlorophyll AB binding protein to produce a transgenic plant, and glucan
 CC is produced in the amyloplast of potato or the vacuole of sugar beet.
 CC Glucans are useful as substitutes for and additions to modified starch
 CC and latexes in paper manufacture. Unlike prior art techniques, which
 CC require input materials that produce chemical effluents, paper
 CC manufacture utilizing the glucan produced by GTF, which utilizes
 CC biologically produced input materials, is more cost-effective and
 CC environmentally friendly. Moreover, glucans also exhibit thermoplastic
 CC properties and impart gloss to the paper during coating step.
 CC The present sequence represents a GTFB mutant of the invention.
 CC Note: The present sequence is not shown in the specification but
 CC was created by the indexer using the GTFB sequence appearing as AAU98027
 CC and the information in claim 36.
 XX
 SQ Sequence 1475 AA;
 XX
 Query Match 100.0%; Score 112; DB 23; Length 1475;
 Best Local Similarity 100.0%; Pred. No. 3.8e-10;
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 TGARTINGQLLYPRANGVQVKG 22
 Db 1300 TGARTINGQLLYPRANGVQVKG 1321
 RESULT 7
 AAU98035
 ID AAU98035 standard; Protein: 1475 AA.
 AC
 AAU98035;
 XX
 27-AUG-2002 (first entry)
 DT
 XX
 S. mutans glucosyltransferase GTFB mutant D457N/D571K.
 DB
 XX
 Glucosyltransferase; GTFB; transgenic plant; paper sizing;
 KM coating composition; glucan; starch; latex; thermoplastic molecule;
 KW amyloplast; vacuole; paper manufacture; mutant; mutain.
 XX
 OS Streptococcus mutans.
 OS Synthetic.
 XX
 Key Location/Qualifiers
 FH Misc-difference 457
 FT /note= "Wild-type Asp substituted by Asn"
 FT Misc-difference 571 /note= "Wild-type Asp substituted by Lys"
 FT
 FT
 XX US2002031826-A1.
 XX
 PD 14-MAR-2002.
 XX
 19-DEC-2000; 2000US-0740274.
 XX
 11-DEC-1998; 98US-0210361.
 PR 07-JUN-1995; 95US-0478704.
 PR 07-JUN-1995; 95US-0482711.
 PR 07-JUN-1995; 95US-0485243.

CC require input materials that produce chemical effluents, paper
 CC manufacture utilizing the glucan produced by GTF, which utilizes
 CC biologically produced input materials, is more cost-effective and
 CC environmentally friendly. Moreover, glucans also exhibit thermoplastic
 CC properties and impact gloss to the paper during coating step.
 CC The present sequence represents a GTFB mutant of the invention.
 CC Note: The present sequence is not shown in the specification but
 CC was created by the indexer using the GTFB sequence appearing as AAU98027
 CC and the information in claim 36.
 XX
 SQ Sequence 1475 AA;
 Query Match 100.0%; Score 112; DB 23; Length 1475;
 Best Local Similarity 100.0%; Pred. No. 3.8e-10;
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 TGARTINGQLLYFRANGVQYKG 22
 Db 1300 TGARTINGQLLYFRANGVQYKG 1321
 RESULT 9
 AAU98037
 ID AAU98037 standard; Protein: 1475 AA.
 AC AAU98037;
 DT 27-AUG-2002 (first entry)
 XX
 DE S. mutans glucosyltransferase GTFB mutant D567T/D571K/K1014T.
 XX
 KW Glucosyltransferase; GTFB; transgenic plant; paper sizing;
 KW coating composition; glucan; starch; latex; thermoplastic molecule;
 KW amyloplast; vacuole; paper manufacture; mutant; mutain.
 XX
 OS Streptococcus mutans.
 OS Synthetic.
 XX
 FH Key Location/Qualifiers
 FT Misc-difference 567
 FT /note= "Wild-type Asp substituted by Thr"
 FT Misc-difference 571
 FT /note= "Wild-type Asp substituted by Lys"
 FT Misc-difference 1014
 FT /note= "Wild-type Lys substituted by Thr"
 XX
 PN US2002031826-A1.
 XX
 PD 14-MAR-2002.
 XX
 PF 19-DEC-2000; 2000US-0740274.
 XX
 PR 11-DEC-1998; 98US-0210361.
 PR 07-JUN-1995; 95US-0478704.
 PR 07-JUN-1995; 95US-0482711.
 PR 07-JUN-1995; 95US-0485243.
 PR 16-JAN-1998; 98US-0007999.
 PR 16-JAN-1998; 98US-0008172.
 PR 20-JAN-1998; 98US-0009620.
 XX
 PA (NICH/) NICHOLS S E.
 XX
 PI Nichols SE;
 XX
 DR WPI; 2002-414332/44.
 XX
 PT Glucosyltransferase B or D protein useful for producing a glucan useful
 PT as substitutes for and additions to modified starch and latexes in
 PT paper manufacture, comprises mutations in specific positions -
 PS Claim 36; Page -: 44pp; English.
 XX
 CC The invention an isolated protein comprising a glucosyltransferase

CC (GTF) B polypeptide having changes at position from I448V, D457N,
 CC D567T, K1014T, D457N/D567T, D457N/D571K, D567T/D571K,
 CC D567T/D571K/K1014T, I448V/D457N/D567T/D571K/K779Q/K1014T,
 CC Y169A/Y170A/Y171A, and K779Q or a GTF D polypeptide having
 CC changes at positions from T589D, T589E, N471D, N471D/T589D, and
 CC N471D/T589E. Also included are a glucan produced by the GTF mutant,
 CC an isolated polynucleotide which encodes P1 or P2, or its complementary
 CC polynucleotide, a ribonucleic acid sequence encoding the GTF mutant,
 CC an expression cassette comprising the polynucleotide operably linked to a
 CC promoter, a vector comprising the expression cassette, host cell
 CC introduced with the vector, a transgenic plant comprising the
 CC vector, a seed or tuber from the transgenic plant, a paper sizing and/or
 CC coating composition comprising a glucan produced in a plant transformed
 CC with a gene encoding the mutant GTF, wild type or, starch, a latex,
 CC thermoplastic molecule or their combinations or glucan and starch where
 CC the glucan is produced in the amyloplast and/or vacuole or a maize line
 CC deficient in starch biosynthesis, transformed with a gene encoding a
 CC glucosyltransferase B or D enzyme, wild-type or mutant and a paper
 CC comprising the glucan (paper sizing/coating agent). The vector is useful
 CC for producing a glucan in a plant. The method comprises transforming a
 CC plant cell with the vector, growing the plant cell under plant growing
 CC conditions to produce a regenerated plant and inducing expression of the
 CC polynucleotide for a time sufficient to produce the glucan in the
 CC regenerated plant, where the vector contains a transit sequence from
 CC ribulose biphosphate carboxylase small subunit, waxy, brittle-1 and
 CC chlorophyll AB binding protein to produce a transgenic plant, and glucan
 CC is produced in the amyloplast of potato or the vacuole of sugar beet.
 CC Glucans are useful as substitutes for and additions to modified starch
 CC and latexes in paper manufacture. Unlike prior art techniques, which
 CC require input materials that produce chemical effluents, paper
 CC manufacture utilizing the glucan produced by GTF, which utilizes
 CC biologically produced input materials, is more cost-effective and
 CC environmentally friendly. Moreover, glucans also exhibit thermoplastic
 CC properties and impact gloss to the paper during coating step.
 CC The present sequence represents a GTFB mutant of the invention.
 CC Note: The present sequence is not shown in the specification but
 CC was created by the indexer using the GTFB sequence appearing as AAU98027
 CC and the information in claim 36.
 XX
 SQ Sequence 1475 AA;
 Query Match 100.0%; Score 112; DB 23; Length 1475;
 Best Local Similarity 100.0%; Pred. No. 3.8e-10;
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 TGARTINGQLLYFRANGVQYKG 22
 Db 1300 TGARTINGQLLYFRANGVQYKG 1321
 RESULT 10
 AAU98038
 ID AAU98038 standard; Protein: 1475 AA.
 AC AAU98038;
 DT 27-AUG-2002 (first entry)
 XX
 DE S. mutans GTFB mutant I448V/D457N/D567T/D571K/K779Q/K1014T.
 XX
 KW Glucosyltransferase; GTFB; transgenic plant; paper sizing;
 KW coating composition; glucan; starch; latex; thermoplastic molecule;
 KW amyloplast; vacuole; paper manufacture; mutant; mutain.
 XX
 OS Streptococcus mutans.
 OS Synthetic.
 XX
 FH Key Location/Qualifiers
 FT Misc-difference 448
 FT /note= "Wild-type Ile substituted by Val"
 FT Misc-difference 457
 FT /note= "Wild-type Asp substituted by Asn"
 FT Misc-difference 567

FT /note= "Wild-type Asp substituted by Thr"
 FT Misc-difference 571
 FT /note= "Wild-type Asp substituted by Lys"
 FT Misc-difference 779
 FT /note= "Wild-type Lys substituted by Gln"
 FT Misc-difference 1014
 FT /note= "Wild-type Lys substituted by Thr"
 FT US2002031826-A1.
 PD 14-MAR-2002.
 PD 19-DEC-2000; 2000US-0740274.
 XX 11-DEC-1998; 98US-0210361.
 PR 07-JUN-1995; 95US-0478704.
 PR 07-JUN-1995; 95US-0482711.
 PR 07-JUN-1995; 95US-0485243.
 PR 16-JAN-1998; 98US-0007999.
 PR 16-JAN-1998; 98US-0008172.
 PR 20-JAN-1998; 98US-0009620.
 XX (NICH/) NICHOLS S E.
 PA Nichols SE;
 PI WPI; 2002-414332/44.
 DR Glucosyltransferase B or D protein useful for producing a glucan useful
 XX as substitutes for and additions to modified starch and latexes in
 PT paper manufacture, comprises mutations in specific positions -
 PT Claim 36; Page -; 44pp; English.
 XX
 CC The invention an isolated protein comprising a glucosyltransferase
 CC (GTF) B polypeptide having changes at position from 1448V, D457N,
 CC D567T, K1014T, D457N/D567T, D457N/D571K, D567T/D571K,
 CC D567T/D571K/K1014T, 1448V/D457N/D567T/K779Q/K1014T,
 CC Y169A/Y170A/Y171A, and K779Q or a GTF D polypeptide having
 CC changes at positions from T589D, T589E, N471D, N471D/T589D, and
 CC N471D/T589E. Also included are a glucan produced by the GTF mutant,
 CC an isolated polynucleotide which encodes P1 or P2, or its complementary
 CC polynucleotide, a ribonucleic acid sequence encoding the GTF mutant,
 CC an expression cassette comprising the polynucleotide operably linked to a
 CC promoter, a vector comprising the expression cassette, host cell
 CC introduced with the vector, a transgenic plant comprising the
 CC vector, a seed or tuber from the transgenic plant, a paper sizing and/or
 CC coating composition comprising a glucan produced in a plant transformed
 CC with a gene encoding the mutant GTF, wild type or, starch, a latex,
 CC thermoplastic molecule or their combinations or glucan and starch where
 CC the glucan is produced in the amyloplast and/or vacuole or a maize line
 CC deficient in starch biosynthesis, transformed with a gene encoding a
 CC glucosyltransferase B or D enzyme, wild-type or mutant and a paper
 CC comprising the glucan (paper sizing/coating agent). The vector is useful
 CC for producing a glucan in a plant. The method comprises transforming a
 CC plant cell with the vector, growing the plant cell under plant growing
 CC conditions to produce a regenerated plant and inducing expression of the
 CC polynucleotide for a time sufficient to produce the glucan in the
 CC regenerated plant, where the vector contains a transit sequence from
 CC ribulose biphosphate carboxylase small subunit, waxy, D11E1-1 and
 CC chloroplast AB binding protein to produce a transgenic plant, and glucan
 CC is produced in the amyloplast of potato or the vacuole of sugar beet.
 CC Glucans are useful as substitutes for and additions to modified starch
 CC and latexes in paper manufacture. Unlike prior art techniques, which
 CC require input materials that produce chemical effluents, paper
 CC manufacture utilizing the glucan produced by GTF, which utilizes
 CC biologically produced input materials, is more cost-effective and
 CC environmentally friendly. Moreover, glucans also exhibit thermoplastic
 CC properties and impart gloss to the paper during coating step.
 CC The present sequence represents a GTFB mutant of the invention.
 CC Note: The present sequence is not shown in the specification but
 CC was created by the indexer using the GTFB sequence appearing as AAU98027
 CC and the information in claim 36.

XX SQ Sequence 1475 AA;
 Query Match 100.0%; Score 112; DB 23; Length 1475;
 Best Local Similarity 100.0%; Pred. No. 3.8e-10;
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Cy 1 TGARTINGQLLYFRANGVQVKG 22
 Db 1300 TGARTINGQLLYFRANGVQVKG 1321
 RESULT 11
 ID AAU98039 standard; Protein; 1475 AA.
 XX AAU98039;
 AC 27-AUG-2002 (first entry)
 DT S. mutans glucosyltransferase GTFB mutant YVY169-171AAA.
 XX
 DE Glucosyltransferase; GTFB; transgenic plant; paper sizing;
 KM coating composition; glucan; starch; latex; thermoplastic molecule;
 KM amyloplast; vacuole; paper manufacture; mutant; mutain.
 XX Streptococcus mutans.
 OS Synthetic.
 OS
 FH Key Location/Qualifiers
 FT Misc-difference 169..171
 FT /note= "Wild-type Tyr-Tyr substituted by
 FT Ala-Ala-Ala"
 XX
 XX US2002031826-A1.
 XX 14-MAR-2002.
 PD 19-DEC-2000; 2000US-0740274.
 XX 11-DEC-1998; 98US-0210361.
 PR 07-JUN-1995; 95US-0478704.
 PR 07-JUN-1995; 95US-0482711.
 PR 07-JUN-1995; 95US-0485243.
 PR 16-JAN-1998; 98US-0007999.
 PR 16-JAN-1998; 98US-0008172.
 PR 20-JAN-1998; 98US-0009620.
 XX (NICH/) NICHOLS S E.
 PA Nichols SE;
 PI WPI; 2002-414332/44.
 DR Glucosyltransferase B or D protein useful for producing a glucan useful
 XX as substitutes for and additions to modified starch and latexes in
 PT paper manufacture, comprises mutations in specific positions -
 PT Claim 36; Page -; 44pp; English.
 XX
 CC The invention an isolated protein comprising a glucosyltransferase
 CC (GTF) B polypeptide having changes at position from 1448V, D457N,
 CC D567T, K1014T, D457N/D567T, D457N/D571K, D567T/D571K,
 CC D567T/D571K/K1014T, 1448V/D457N/D567T/K779Q/K1014T,
 CC Y169A/Y170A/Y171A, and K779Q or a GTF D polypeptide having
 CC changes at positions from T589D, T589E, N471D, N471D/T589D, and
 CC N471D/T589E. Also included are a glucan produced by the GTF mutant,
 CC an isolated polynucleotide which encodes P1 or P2, or its complementary
 CC polynucleotide, a ribonucleic acid sequence encoding the GTF mutant,
 CC an expression cassette comprising the polynucleotide operably linked to a
 CC promoter, a vector comprising the expression cassette, host cell
 CC introduced with the vector, a transgenic plant comprising the
 CC vector, a seed or tuber from the transgenic plant, a paper sizing and/or

coating composition comprising a glucan produced in a plant transformed with a gene encoding the mutant GTF, wild type or starch, a latex, thermoplastic molecule or their combinations or glucan and starch where the glucan is produced in the amyloplast and/or vacuole or a maize line deficient in starch biosynthesis, transformed with a gene encoding a glucosyltransferase B or D enzyme, wild-type or mutant and a paper comprising the glucan (paper sizing/coating agent). The vector is useful for producing a glucan in a plant. The method comprises transforming a plant cell with the vector, growing the plant cell under plant growing conditions to produce a regenerated plant and inducing expression of the polynucleotide for a time sufficient to produce the glucan in the regenerated plant, where the vector contains a transit sequence from ribulose biphosphate carboxylase small subunit, waxy, brittle-1 and chlorophyll AB binding protein to produce a transgenic plant, and glucan is produced in the amyloplast of potato or the vacuole of sugar beet. Glucans are useful as substitutes for and additions to modified starch and latexes in paper manufacture. Unlike prior art techniques, which require input materials that produce chemical effluents, paper manufacture utilizing the glucan produced by GTF, which utilizes biologically produced input materials, is more cost-effective and environmentally friendly. Moreover, glucans also exhibit thermoplastic properties and impart gloss to the paper during coating step. The present sequence represents a GTFB mutant of the invention. Note: The present sequence is not shown in the specification but was created by the indexer using the GTFB sequence appearing as AAU98027 and the information in claim 36.

Sequence 1475 AA;

Query Match 100.0%; Score 112; DB 23; Length 1475;
Best Local Similarity 100.0%; Pred. No. 3.8e-10;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TGARTINGQLLYFRANGVQVKG 22
1300 TGARTINGQLLYFRANGVQVKG 1321

RESULT 12
AAU98040

ID AAU98040 standard; Protein; 1475 AA.

AC AAU98040;

DT 27-AUG-2002 (first entry)

DE S. mutans glucosyltransferase GTFB mutant K779Q.

KM Glucosyltransferase; GTFB; transgenic plant; paper sizing;
KM coating composition; glucan; starch; latex; thermoplastic molecule;
KM amyloplast; vacuole; paper manufacture; mutant; mutain.

XX Streptococcus mutans.

OS Synthetic.

EH Key Location/Qualifiers

FT Misc-difference 779 /note= "Wild-type Lys substituted by Gln"

PN US2002031826-A1.

XX 14-MAR-2002.

PF 19-DEC-2000; 2000US-0740274.

XX 11-DEC-1998; 98US-0210361.

PR 07-JUN-1995; 95US-0478704.

PR 07-JUN-1995; 95US-0482711.

PR 16-JAN-1998; 98US-0007939.

PR 20-JAN-1998; 98US-0009620.

PA (NICH/) NICHOLS S E.

XX Nichols SE;

DR WPI: 2002-414332/44.

PT Glucosyltransferase B or D protein useful for producing a glucan useful
PT as substitutes for and additions to modified starch and latexes in
PT paper manufacture, comprises mutations in specific positions -
XX Claim 36; Page -; 44pp; English.

The invention an isolated protein comprising a glucosyltransferase (GTF) B polypeptide having changes at position from 1448V, D457N, D567T, K1014T, D457N/D567T, D457N/D571K, D567T/D571K, D567T/D571K/K1014T, 1448V/D457N/D567T/D571K/K779Q/K1014T, Y159A/Y170A/Y171A, and K779Q or a GTF D polypeptide having changes at positions from T589D, T589E, N471D, N471D/T589D, and N471D/T589E. Also included are a glucan produced by the GTF mutant, an isolated polynucleotide which encodes P1 or P2, or its complementary polynucleotide, a ribonucleic acid sequence encoding the GTF mutant, an expression cassette comprising the polynucleotide operably linked to a promoter, a vector comprising the expression cassette, host cell introduced with the vector, a transgenic plant comprising the vector, a seed or tuber from the transgenic plant, a paper sizing and/or coating composition comprising a glucan produced in a plant transformed with a gene encoding the mutant GTF, wild type or starch, a latex, thermoplastic molecule or their combinations or glucan and starch where the glucan is produced in the amyloplast and/or vacuole or a maize line deficient in starch biosynthesis, transformed with a gene encoding a glucosyltransferase B or D enzyme, wild-type or mutant and a paper comprising the glucan (paper sizing/coating agent). The vector is useful for producing a glucan in a plant. The method comprises transforming a plant cell with the vector, growing the plant cell under plant growing conditions to produce a regenerated plant and inducing expression of the polynucleotide for a time sufficient to produce the glucan in the regenerated plant, where the vector contains a transit sequence from ribulose biphosphate carboxylase small subunit, waxy, brittle-1 and chlorophyll AB binding protein to produce a transgenic plant, and glucan is produced in the amyloplast of potato or the vacuole of sugar beet. Glucans are useful as substitutes for and additions to modified starch and latexes in paper manufacture. Unlike prior art techniques, which require input materials that produce chemical effluents, paper manufacture utilizing the glucan produced by GTF, which utilizes biologically produced input materials, is more cost-effective and environmentally friendly. Moreover, glucans also exhibit thermoplastic properties and impart gloss to the paper during coating step. The present sequence represents a GTFB mutant of the invention. Note: The present sequence is not shown in the specification but was created by the indexer using the GTFB sequence appearing as AAU98027 and the information in claim 36.

Sequence 1475 AA;

Query Match 100.0%; Score 112; DB 23; Length 1475;
Best Local Similarity 100.0%; Pred. No. 3.8e-10;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TGARTINGQLLYFRANGVQVKG 22
1300 TGARTINGQLLYFRANGVQVKG 1321

RESULT 13
AAU79286

ID AAU79286 standard; Protein; 380 AA.

AC AAU79286;

DT 13-AUG-2002 (first entry)

DE Streptococcus mutans monoclonal antibody-related protein #3.

KW Antibody; dental caries; water insoluble glucan synthetase;
 XX anti-carries; glucosyl transferase-B; immunotherapy.
 OS Streptococcus mutans.
 XX JP2002114709-A.
 XX 16-APR-2002.
 XX 04-OCT-2000; 2000JP-0304889.
 XX 04-OCT-2000; 2000JP-0304889.
 XX (UYNI-) UNIV NIPPON.
 XX WPI; 2002-448101/48.
 DR Anti-carries agent composed of a monoclonal antibody against an
 PT inhibitory enzyme against water insoluble glucan synthetase of glucosyl
 PT transferase-B (GTF-B) of Streptococcus mutans -
 XX
 PS Claim 5; Page 19-20; 28pp; Japanese.
 XX The invention relates to a monoclonal antibody against dental caries and
 CC an anti-carries agent composed of a monoclonal antibody produced by
 CC Streptococcus mutans, particularly mouse-hybridoma MHP126 (FERM P-17566)
 CC or mouse-hybridoma MHP136 (FERM P-17567), against an enzyme having
 CC inhibitive activity against water insoluble glucan synthetase of glucosyl
 CC transferase-B. The monoclonal antibody specifically inhibits water
 CC insoluble glucan synthetase of Streptococcus mutans produced glucosyl
 CC transferase-B and is used in the immunotherapy of dental caries. This
 CC sequence represents a Streptococcus mutans monoclonal antibody-related
 CC protein.
 CC
 SQ Sequence 380 AA;
 Query Match 93.8%; Score 105; DB 23; Length 380;
 Best Local Similarity 95.5%; Pred. No. 1.2e-09;
 Matches 21; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 TGARTINGQLTYFRANGVQVKG 22
 DB 205 TGARTINGQLTYFRANGVQVKG 226
 RESULT 14
 AAU79284
 ID AAU79284 standard; Protein; 1476 AA.
 XX AAU79284;
 AC 13-AUG-2002 (first entry)
 DT Streptococcus mutans monoclonal antibody-related protein #1.
 XX Streptococcus mutans monoclonal antibody-related protein #1.
 XX Antibody; dental caries; water insoluble glucan synthetase;
 KW anti-carries; glucosyl transferase-B; immunotherapy.
 XX Streptococcus mutans.
 OS JP2002114709-A.
 XX 16-APR-2002.
 XX 04-OCT-2000; 2000JP-0304889.
 XX 04-OCT-2000; 2000JP-0304889.
 XX (UYNI-) UNIV NIPPON.
 XX WPI; 2002-448101/48.
 DR Anti-carries agent composed of a monoclonal antibody against an

PT inhibitory enzyme against water insoluble glucan synthetase of glucosyl
 PT transferase-B (GTF-B) of Streptococcus mutans -
 XX
 PS Claim 3; Page 13-16; 28pp; Japanese.
 XX The invention relates to a monoclonal antibody against dental caries and
 CC an anti-carries agent composed of a monoclonal antibody produced by
 CC Streptococcus mutans, particularly mouse-hybridoma MHP126 (FERM P-17566)
 CC or mouse-hybridoma MHP136 (FERM P-17567), against an enzyme having
 CC inhibitive activity against water insoluble glucan synthetase of glucosyl
 CC transferase-B. The monoclonal antibody specifically inhibits water
 CC insoluble glucan synthetase of Streptococcus mutans produced glucosyl
 CC transferase-B and is used in the immunotherapy of dental caries. This
 CC sequence represents a Streptococcus mutans monoclonal antibody-related
 CC protein.
 CC
 SQ Sequence 1476 AA;
 Query Match 93.8%; Score 105; DB 23; Length 1476;
 Best Local Similarity 95.5%; Pred. No. 6.3e-09;
 Matches 21; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 TGARTINGQLTYFRANGVQVKG 22
 DB 1301 TGARTINGQLTYFRANGVQVKG 1322
 RESULT 15
 ID AAR43696 standard; peptide; 22 AA.
 XX AAR43696;
 AC 25-MAR-2003 (updated)
 DT 20-MAY-1994 (first entry)
 XX GT subsequence.
 DE GT; glucosyltransferase; vaccine; T-cell; B-cell; reaction;
 KW immunoreponse; peptidyl core matrix; dental caries; diptheria;
 KW tetanus; measles; polio.
 XX
 OS Synthetic.
 XX WO9322341-A1.
 XX 11-NOV-1993.
 PD 30-APR-1993; 93WO-US04094.
 XX 01-MAY-1992; 92US-0877295.
 XX (FORS-) FORSYTH DENTAL INFIRMARY FOR CHILDREN.
 XX Smith DJ, Taubman MA;
 DR WPI; 1993-368721/46.
 XX Synthetic glucosyl-transferase peptide(s) - provoke T- and B-cell
 PT reactions in mammals, and are effective against dental caries
 XX
 PS Claim 3; Page 23; 38pp; English.
 XX The sequences (AAR43694-98) are subsequences from GT, they provoke
 CC T- and B-cell responses in mammals. The can be used to produce a
 CC vaccine comprising 2 of these sequences attached to a peptidyl
 CC core matrix. They are useful in producing T-cell responses to
 CC the GT enzyme causing a reduction of colonisation or accumulation
 CC of mutans streptococcal strains in mammals. The vaccines can be used
 CC in preventing dental caries.
 CC (Updated on 25-MAR-2003 to correct PN field.)
 CC
 SQ Sequence 22 AA;

Query Match 77.7%; Score 87; DB 14; Length 22;
 Best Local Similarity 77.3%; Pred. No. 4.6e-08;
 Matches 17; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 TGARTINGOLLYPRANGOVKG 22
 |||:|||||
 Db 1 TGAQTIKGQKLYFKANGQOVKG 22

Search completed: November 13, 2003, 09:38:26
 Job time : 41.3507 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 13, 2003, 09:31:40, Search time 13.6588 Seconds
(without alignments)
154.898 Million cell updates/sec

Title: US-09-290-049A-3

Sequence: 1 TGARTINGQLLYFRANGVQVKG 22

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: PIR 76:*

1: PIR1:.*
2: PIR2:.*
3: PIR3:.*
4: PIR4:.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	112	100.0	1475	2	B3135
2	87	77.7	1592	2	A38175
3	82	73.2	1599	2	S22737
4	80	71.4	1518	2	A44811
5	78	69.6	1375	2	UT0345
6	75	67.0	1290	2	UC5473
7	73	65.2	1365	2	A41483
8	72	64.3	1449	2	T30857
9	72	64.3	1508	2	T31098
10	68	60.7	563	2	A37184
11	68	60.7	1449	2	T30552
12	67	59.8	1431	2	A45866
13	66	58.9	1577	2	T30858
14	65.5	58.5	329	2	A55221
15	55	49.1	2817	2	B97033
16	54	48.2	221	2	T37244
17	50	44.6	831	2	T00323
18	49.5	44.2	529	2	T10388
19	49.5	44.2	2710	2	A37052
20	49	43.8	491	2	C97069
21	47.5	42.4	240	2	F97067
22	47	42.0	126	2	S40016
23	47	42.0	696	2	A28635
24	47	42.0	788	2	US0747
25	46	41.1	231	2	F70471
26	46	41.1	245	2	F64465
27	46	41.1	301	2	S13165
28	46	41.1	619	2	A41971
29	46	41.1	619	2	A97887

30	45	41.1	744	2	F95013	pneumococcal surfa
31	46	41.1	1301	2	S18118	alpha-amylase - Al
32	46	41.1	2364	2	T40884	cytochrome b - C10
33	45	40.2	142	2	F87265	conserved hypotet
34	45	40.2	630	2	T05433	hypothetical prote
35	45	40.2	648	2	S10869	enterotoxin A - C1
36	44	39.3	158	2	A12117	hypothetical prote
37	44	39.3	329	2	F71014	probable dipeptid
38	44	39.3	522	2	D71074	probable methylal
39	44	39.3	522	2	H75135	methylmalonyl-coa
40	44	39.3	566	2	S74633	high affinity sulf
41	44	39.3	591	2	C84220	propionyl-CoA carb
42	44	39.3	601	2	AG0066	probable AMP-bind
43	44	39.3	829	2	T29372	hypothetical prote
44	44	39.3	950	2	B87611	TonB-dependent rec
45	44	39.3	1396	2	S36851	L-shaped tail fibe

ALIGNMENTS

```
RESULT 1
B3135
gluB protein precursor - Streptococcus mutans
C:Species: Streptococcus mutans
C:Date: 23-Oct-1990 #sequence_revision 23-Oct-1990 #text_change 15-Oct-1999
C:Accession: B3135, A33128
R:Shiroza, T.; Ueda, S.; Kuramitsu, H. K.
J. Bacteriol. 169, 4263-4270, 1987
A:Title: Sequence analysis of the gluB gene from Streptococcus mutans.
A:Reference number: A33135, MUID:87308013, PMID:3040685
A:Accession: B3135
A:Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-1475 <SH1>
A:Cross-references: CB:M17361; NID:G153639; PIDN:AA88588.1; PID:G153640
R:Shiroza, T.; Ueda, S.; Kuramitsu, H. K.
Submitted to the Protein Sequence Database, September 1990
A:Reference number: A33128
A:Accession: A33128
A:Status: Preliminary; not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 1-171,173-641,'N',643-1475 <SH2>
A:Experimental source: strain GS-5
C:Superfamily: cpl repeat homology
F:1096-1115/Domain: cpl repeat homology <CP1>
F:1224-1243/Domain: cpl repeat homology <CP2>
F:1289-1308/Domain: cpl repeat homology <CP3>
F:1354-1373/Domain: cpl repeat homology <CP4>
F:1419-1438/Domain: cpl repeat homology <CP5>

Query Match 100.0%; Score 112; DB 2; Length 1475;
Best Local Similarity 100.0%; Pred. No. 6.1e-10;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

CY 1 TGARTINGQLLYFRANGVQVKG 22
DB 1300 TGARTINGQLLYFRANGVQVKG 1321

RESULT 2
A38175
glucosyltransferase precursor - Streptococcus sobrinus
C:Species: Streptococcus sobrinus
C:Date: 28-Aug-1992 #sequence_revision 28-Aug-1992 #text_change 15-Oct-1999
C:Accession: A38175
R:Abu, H.; Mtsunura, T.; Kodama, T.; Ohta, H.; Fukui, K.; Kato, K.; Kagawa, H.
J. Bacteriol. 173, 989-996, 1991
A:Title: Peptide sequences for sucrose splitting and glucan binding within Streptococcus
A:Reference number: A38175; MUID:91123227; PMID:1704006
A:Accession: A38175
A:Status: Preliminary
A:Molecule type: DNA
```

A/Residues: 1-1592 <ABO>
 A/Cross-references: GB:D90213; NID:9217032; PIDN:BAAL241.1; PID:dl014946; PID:9217033
 C/Superfamily: cpl repeat homology
 F:1093-1112/Domain: cpl repeat homology <CP1>
 F:1222-1241/Domain: cpl repeat homology <CP2>
 F:1287-1306/Domain: cpl repeat homology <CP3>
 F:1330-1351/Domain: cpl repeat homology <CP4>
 F:1352-1371/Domain: cpl repeat homology <CP5>
 F:1402-1420/Domain: cpl repeat homology <CP6>
 F:1465-1484/Domain: cpl repeat homology <CP7>
 F:1513-1532/Domain: cpl repeat homology <CP8>

Query Match 77.7%; Score 87; DB 2; Length 1592;
 Best Local Similarity 77.3%; Pred. No. 8.9e-06;
 Matches 17; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 TGARTINGQLLYFRANGVQYKG 22
 DB 1298 TGAQITRGQKLYFKANGQYKG 1319

RESULT 3
 S22737
 glucosyltransferase (EC 2.4.1.-) S - Streptococcus salivarius
 C/Species: Streptococcus salivarius
 C/Date: 16-Sep-1992 #sequence_revision 16-Sep-1992 #text_change 21-Jan-2000
 C/Accession: S22737; S28810; B44811; S22727
 R/Jacques, N. Submitted to the EMBL Data Library, March 1992
 A/Reference number: S22726
 A/Accession: S22737
 A/Molecule type: DNA
 A/Residues: 1-1599 <JAC>
 A/Cross-references: EMBL:Z11872; NID:g47530; PIDN:CAA77698.1; PID:g47531
 A/Experimental source: ATCC 29975
 R/Giffard, P.M.; Simpson, C.L.; Milward, C.P.; Jacques, N.A.
 J. Gen. Microbiol. 137, 2577-2593, 1991
 A/Title: Molecular characterization of a cluster of at least two glucosyltransferase gen
 A/Reference number: A44811; MUID:92148377; PMID:1838391
 A/Accession: S28810
 A/Molecule type: DNA
 A/Residues: 1-51 <GIF>
 A/Cross-references: EMBL:Z11873
 C/Genetics: gtfK
 A/Gene: gtfK
 C/Superfamily: cpl repeat homology
 C/Keywords: glucosyltransferase; hexosyltransferase
 F:1456-1475/Domain: cpl repeat homology <CPR>

Query Match 73.2%; Score 82; DB 2; Length 1599;
 Best Local Similarity 77.3%; Pred. No. 6e-05;
 Matches 17; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 TGARTINGQLLYFRANGVQYKG 22
 DB 1402 TGAQVINGQLLYFDANGRQYKG 1423

RESULT 4
 A44811
 glucosyltransferase (EC 2.4.1.-) I - Streptococcus salivarius
 C/Species: Streptococcus salivarius
 C/Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 15-Oct-1999
 C/Accession: A44811; S22726; S28809
 R/Giffard, P.M.; Simpson, C.L.; Milward, C.P.; Jacques, N.A.
 J. Gen. Microbiol. 137, 2577-2593, 1991
 A/Title: Molecular characterization of a cluster of at least two glucosyltransferase ge
 A/Reference number: A44811; MUID:92148377; PMID:1838391
 A/Accession: A44811
 A/Molecule type: DNA
 A/Residues: 1-1518 <GIF>
 A/Cross-references: EMBL:Z11873; NID:g47526; PIDN:CAA77900.1; PID:g47527
 A/Note: sequence extracted from NCBI backbone (NCBIT:81050, NCSTP:81052)

C/Genetics: gtfJ
 A/Gene: gtfJ
 C/Superfamily: cpl repeat homology
 C/Keywords: glucosyltransferase; hexosyltransferase
 F:1307-1326/Domain: cpl repeat homology <CP4>

Query Match 71.4%; Score 80; DB 2; Length 1518;
 Best Local Similarity 72.7%; Pred. No. 0.00012;
 Matches 16; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 TGARTINGQLLYFRANGVQYKG 22
 DB 1383 TGAQVINGQLLYFRANGVQYKG 1404

RESULT 5
 J0345
 dextranucrase (EC 2.4.1.5) precursor - Streptococcus mutans (strain GS-5)
 N/Alternate names: sucrose 6-glucosyltransferase
 C/Species: Streptococcus mutans
 C/Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 24-Sep-1999
 C/Accession: J0345; C33135
 R/Ueda, S.; Shiroza, T.; Kuramitsu, H.K.
 Gene 69, 101-109, 1988
 A/Title: Sequence analysis of the gtfC gene from Streptococcus mutans GS-5.
 A/Reference number: J0345; MUID:89137980; PMID:2976010
 A/Accession: J0345
 A/Molecule type: DNA
 A/Residues: 1-1375 <UED>
 A/Experimental source: GS-5
 R/Shiroza, T.; Ueda, S.; Kuramitsu, H.K.
 J. Bacteriol. 169, 4263-4270, 1987
 A/Title: Sequence analysis of the gtfB gene from Streptococcus mutans.
 A/Reference number: A33135; MUID:87308013; PMID:3040685
 A/Accession: C33135
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-349 <SHI>
 A/Cross-references: GB:M17361
 C/Genetics: gtfC
 A/Gene: gtfC
 C/Punctuation:
 A/Description: catalyzes the synthesis of both water-soluble and water-insoluble glucans
 C/Superfamily: cpl repeat homology
 C/Keywords: duplication; glycosyltransferase; hexosyltransferase
 F:1-34/Domain: signal sequence #status predicted <SIG>
 F:135-1375/Product: glucosyltransferase #status predicted <MAT>
 F:1126-1145/Domain: cpl repeat homology <CP1>
 F:1253-1272/Domain: cpl repeat homology <CP2>
 F:1318-1337/Domain: cpl repeat homology <CP3>

Query Match 69.6%; Score 78; DB 2; Length 1375;
 Best Local Similarity 68.2%; Pred. No. 0.00023;
 Matches 15; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 1 TGARTINGQLLYFRANGVQYKG 22
 DB 1264 TGTVTENGQLLYFKPVGQYKG 1285

RESULT 6
 J03473
 dextranucrase (EC 2.4.1.5) - Leuconostoc mesenteroides
 C/Species: Leuconostoc mesenteroides
 C/Date: 07-Jul-1997 #sequence_revision 29-Aug-1997 #text_change 29-Aug-1997
 C/Accession: J03473
 R/Monchois, V.; Millemot, R.M.; Remaud-Simeon, M.; Croux, C.; Monsan, P.
 Gene 182, 23-32, 1996
 A/Title: Cloning and sequencing of a gene coding for a novel dextranucrase from Leucon
 A/Reference number: J03473; MUID:9713686; PMID:8982063
 A/Accession: J03473
 A/Status: nucleic acid sequence not shown; translation not shown
 A/Molecule type: DNA

A;Residues: 1-1290 <MON>
A;Cross-references: GB:U38181
C;Comment: This enzyme catalyzes the transfer of D-glucopyranosyl units from sucrose onto
C;Genetics:
A;Gene: dera
C;Keywords: glycosyltransferase; hexosyltransferase
F;78-870/Domain: catalytic #status predicted <CAT>
F;922-1290/Domain: glucan-binding #status predicted <GCB>

Query Match 67.0%; Score 75; DB 2; Length 1290;
Best Local Similarity 68.2%; Pred. No. 0.00068;
Matches 15; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 1 TGARTINGQLYFRANGVQVKG 22
Db 1245 TGVQVINGQLYFPADGRQVKG 1266

RESULT 7

A41483
glucosyltransferase (EC 2.4.1.-) gtf5 precursor - Streptococcus sobrinus
C;Species: Streptococcus sobrinus
C;Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 15-Oct-1999
C;Accession: A41483
R;Gilmore, K.S.; Russell, R.R.B.; Ferretti, J.J.
Infected. Immun. 58, 2452-2458, 1990
A;Title: Analysis of the Streptococcus downei gtf5 gene, which specifies a glucosyltrans
A;Reference number: A41483; MUID:90316665; PMID:2142479
A;Accession: A41483
A;Molecule type: DNA
A;Residues: 1-1365 <GIL>
A;Cross-references: GB:M30943; NID:g153652; PIDN:AAA26898.1; PID:g153653
C;Genetics:
A;Gene: gtf5
C;Superfamily: cpl repeat homology
C;Keywords: glycosyltransferase; hexosyltransferase
Query Match 67.0%; Score 75; DB 2; Length 1365;
Best Local Similarity 63.6%; Pred. No. 0.00072;
Matches 14; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

QY 1 TGARTINGQLYFRANGVQVKG 22
Db 1235 TGEQTIDGQKVFQDNGVQVKG 1256

RESULT 8

T30857
glucosyltransferase - Streptococcus salivarius
C;Species: Streptococcus salivarius
C;Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 22-Oct-1999
C;Accession: T30857
R;Stimpson, C.L.; Giffard, P.M.; Jacques, N.A.
Infected. Immun. 63, 609-621, 1995
A;Title: Streptococcus salivarius ATCC 25975 possesses at least two genes coding for pri
A;Reference number: Z20909; MUID:95122197; PMID:7822030
A;Accession: T30857
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-1449 <SIM>
A;Cross-references: EMBL:L35495; NID:g662378; PID:g662379; PIDN:AAC41412.1
C;Genetics:
A;Gene: gtfL

Query Match 65.2%; Score 73; DB 2; Length 1449;
Best Local Similarity 68.2%; Pred. No. 0.0016;
Matches 15; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 1 TGARTINGQLYFRANGVQVKG 22
Db 1258 TGHONINGQLYFRANGVQVKG 1279

RESULT 9
T31098
probable dextranucrase (EC 2.4.1.5), extracellular - Leuconostoc mesenteroides

C;Species: Leuconostoc mesenteroides
C;Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 11-May-2000
C;Accession: T31098
R;Monchoix, V.; Renaud-Simeon, M.; Monsan, P.; Willemot, R.M.
FEMS Microbiol. Lett. 159, 307-315, 1998
A;Title: Cloning and sequencing of a gene coding for an extracellular dextranucrase (DS
A;Reference number: Z20981; MUID:98164374; PMID:9503626
A;Accession: T31098
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-1508 <MON>
A;Cross-references: EMBL:AF030129; NID:g2766611; PID:g2766612; PIDN:AAB95453.1
A;Experimental source: strain NRRL B-1299
C;Genetics:
A;Gene: dera
C;Function:
A;Description: produces dextran composed only of alpha(1-6) glucosidic bonds
C;Keywords: glycosyltransferase; hexosyltransferase

Query Match 64.3%; Score 72; DB 2; Length 1508;
Best Local Similarity 68.2%; Pred. No. 0.0025;
Matches 15; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 1 TGARTINGQLYFRANGVQVKG 22
Db 1399 TGLVNINGMLKTFQANGROVKG 1420

RESULT 10

A37184
glucan-binding protein - Streptococcus mutans
C;Species: Streptococcus mutans
C;Date: 31-Jan-1992 #sequence_revision 31-Jan-1992 #text_change 15-Oct-1999
C;Accession: A37184
R;Banat, J.A.; Russell, R.R.B.; Ferretti, J.J.
Infected. Immun. 58, 667-673, 1990
A;Title: Sequence analysis of the gene for the glucan-binding protein of Streptococcus m
A;Reference number: A37184; MUID:90170123; PMID:2307516
A;Accession: A37184
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-563 <BAN>
A;Cross-references: GB:M30945; NID:g153637; PIDN:AAA26894.1; PID:g153638
C;Superfamily: cpl repeat homology
C;Date: 31-Jan-1992 #sequence_revision 31-Jan-1992 #text_change 15-Oct-1999
F;169-188/Domain: cpl repeat homology <CP1>
F;264-283/Domain: cpl repeat homology <CP2>
F;349-368/Domain: cpl repeat homology <CP3>
F;504-523/Domain: cpl repeat homology <CP4>
F;525-548/Domain: cpl repeat homology <CP5>

Query Match 60.7%; Score 68; DB 2; Length 563;
Best Local Similarity 66.7%; Pred. No. 0.0037;
Matches 14; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 2 GARTINGQLYFRANGVQVKG 22
Db 276 GWRITGKRYTFDNGVQVKG 296

RESULT 11

T30552
glucosyltransferase N - Streptococcus salivarius (fragment)
C;Species: Streptococcus salivarius
C;Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 22-Oct-1999
C;Accession: T30552
R;Jaife, R.I.
submitted to the EMBL Data Library, February 1998
A;Description: Streptococcus salivarius V1477 gtfN.
A;Reference number: Z20854
A;Accession: T30552

A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1449 <UAF>
A:Cross-references: EMBL:AF049609; NID:g2935545; PID:g2935546; PIDN:AAC05156.1
C:Genetics:
A:Gene: gtfM

Query Match 60.7%; Score 68; DB 2; Length 1449;
Best Local Similarity 66.7%; Pred. No. 0.011;
Matches 14; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 2 TGARTINGQLLYFRANGVQVKG 22
DB 1259 GHQNINGDELFEDNNQVQVKG 1279

RESULT 12
A:Accession: A45866
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1431 <HON>
A:Cross-references: GB:M29296
A:Superfamily: cpl repeat homology
C:Keywords: glycosyltransferase; hexosyltransferase
F:181-201/Domain: cpl repeat homology <CP1>
F:1127-1146/Domain: cpl repeat homology <CP2>
F:1192-1211/Domain: cpl repeat homology <CP3>
F:1252-1276/Domain: cpl repeat homology <CP4>
F:1277-1297/Domain: cpl repeat homology <CP5>
F:1321-1340/Domain: cpl repeat homology <CP6>
F:1341-1361/Domain: cpl repeat homology <CP6>
F:1385-1404/Domain: cpl repeat homology <CP7>

A:Title: Nucleotide sequence of the Streptococcus mutans gtfD gene encoding the glucosyl
A:Reference number: A45866; MUID:91100958; PMID:2148600

A:Accession: A45866
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1431 <HON>
A:Cross-references: GB:M29296
A:Superfamily: cpl repeat homology
C:Keywords: glycosyltransferase; hexosyltransferase
F:181-201/Domain: cpl repeat homology <CP1>
F:1127-1146/Domain: cpl repeat homology <CP2>
F:1192-1211/Domain: cpl repeat homology <CP3>
F:1252-1276/Domain: cpl repeat homology <CP4>
F:1277-1297/Domain: cpl repeat homology <CP5>
F:1321-1340/Domain: cpl repeat homology <CP6>
F:1341-1361/Domain: cpl repeat homology <CP6>
F:1385-1404/Domain: cpl repeat homology <CP7>

A:Accession: A45866
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1431 <HON>
A:Cross-references: GB:M29296
A:Superfamily: cpl repeat homology
C:Keywords: glycosyltransferase; hexosyltransferase
F:181-201/Domain: cpl repeat homology <CP1>
F:1127-1146/Domain: cpl repeat homology <CP2>
F:1192-1211/Domain: cpl repeat homology <CP3>
F:1252-1276/Domain: cpl repeat homology <CP4>
F:1277-1297/Domain: cpl repeat homology <CP5>
F:1321-1340/Domain: cpl repeat homology <CP6>
F:1341-1361/Domain: cpl repeat homology <CP6>
F:1385-1404/Domain: cpl repeat homology <CP7>

A:Accession: A45866
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1431 <HON>
A:Cross-references: GB:M29296
A:Superfamily: cpl repeat homology
C:Keywords: glycosyltransferase; hexosyltransferase
F:181-201/Domain: cpl repeat homology <CP1>
F:1127-1146/Domain: cpl repeat homology <CP2>
F:1192-1211/Domain: cpl repeat homology <CP3>
F:1252-1276/Domain: cpl repeat homology <CP4>
F:1277-1297/Domain: cpl repeat homology <CP5>
F:1321-1340/Domain: cpl repeat homology <CP6>
F:1341-1361/Domain: cpl repeat homology <CP6>
F:1385-1404/Domain: cpl repeat homology <CP7>

Query Match 59.8%; Score 67; DB 2; Length 1431;
Best Local Similarity 59.1%; Pred. No. 0.016;
Matches 13; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY 1 TGARTINGQLLYFRANGVQVKG 22
DB 1332 TGSQTLAGKLYFASDQKQVKG 1353

RESULT 13
T30858

glucosyltransferase - Streptococcus salivarius
C:Species: Streptococcus salivarius
C:Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 22-Oct-1999
C:Accession: T30858
R:Simpson, C.L.; Giffard, P.M.; Jacques, N.A.
Infect. Immun. 63, 609-621, 1995

A:Title: Streptococcus salivarius ATCC 25975 possesses at least two genes coding for pri
A:Reference number: Z20909; MUID:95122197; PMID:7822030
A:Accession: T30858
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1577 <SIM>
A:Cross-references: EMBL:J35928; NID:g662380; PID:g662381; PIDN:AAC41413.1
C:Genetics:
A:Gene: gtfM

Query Match 58.9%; Score 66; DB 2; Length 1577;
Best Local Similarity 54.5%; Pred. No. 0.026;
Matches 12; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

QY 1 TGARTINGQLLYFRANGVQVKG 22
DB 1469 TGLQNNKRYFGSNGAQVKG 1490

RESULT 14
A55221

dextranase inhibitor precursor - Streptococcus sobrinus (strain UAB66, serotype g)
C:Species: Streptococcus sobrinus
C:Date: 23-Mar-1995 #sequence_revision 23-Mar-1995 #text_change 17-Mar-1999
C:Accession: A55221
R:Sun, J.W.; Wanda, S.Y.; Camilli, A.; Curtiss III, R.
J. Bacteriol. 176, 7213-7222, 1994

A:Title: Cloning and DNA sequencing of the dextranase inhibitor gene (dei) from Streptoc
A:Reference number: A55221; MUID:95050304; PMID:7961493
A:Accession: A55221
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-329 <SUN>
A:Cross-references: GB:L34406
C:Superfamily: cpl repeat homology
F:173-192/Domain: cpl repeat homology <CP1>

Query Match 58.5%; Score 65.5; DB 2; Length 329;
Best Local Similarity 68.2%; Pred. No. 0.0056;
Matches 15; Conservative 1; Mismatches 5; Indels 1; Gaps 1;

QY 2 GARTINGQLLYFRAN-GVQVKG 22
DB 185 GLQTINGQLLYFDNTGKQVKG 206

RESULT 15

uncharacterized protein, related to enterotoxins of other Clostridiales [imported] - Clo

C:Species: Clostridium acetobutylicum
C:Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 14-Sep-2001

C:Accession: B97033
R:Nolling, J.; Breton, G.; Omeichenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee,
J. Bacteriol. 183, 4823-4838, 2001

A:Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium Clo
A:Reference number: A96900; MUID:21359325; PMID:21359325
A:Accession: B97033
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-2817 <KUR>
A:Cross-references: GB:AE001437; PIDN:AAK79053.1; PID:G15023993; GSPDB:GN00168
A:Experimental source: Clostridium acetobutylicum ATCC824
C:Genetics:
A:Gene: CAC1079

Query Match 49.1%; Score 55; DB 2; Length 2817;
Best Local Similarity 45.5%; Pred. No. 3.2;
Matches 10; Conservative 4; Mismatches 8; Indels 0; Gaps 0;

QY 1 TGARTINGQLLYFRANGVQVKG 22
DB 1469 TGLQNTDNTTYFDNSGIMQVKG 1490

Search completed: November 13, 2003, 09:50:24
Job time: 15.6588 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 13, 2003, 09:31:40 ; Search time 7.32417 Seconds
(without alignments)
130.561 Million cell updates/sec

Title: US-09-290-049a-3

Perfect score: 112
Sequence: 1 TGARTINGQLLYFRANGVQVKG 22

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	105	93.8	1476	1	GTFB_STRMU P08987 streptococc
2	87	77.7	1592	1	GTF2_STRDO P27470 streptococc
3	87	77.7	1597	1	GTF1_STRDO P11001 streptococc
4	86	76.8	1455	1	GTF3_STRMU P13470 streptococc
5	75	67.0	1365	1	GTF5_STRDO P29336 streptococc
6	67	59.8	1462	1	GTFD_STRMU P49331 streptococc
7	49.5	44.2	529	1	Y119_NPYOP C10358 oxygia pseu
8	49.5	44.2	2710	1	TOXA_CLODI P16154 clostridium
9	47.5	42.4	240	1	BIOD_CLOAB Q97JCS clostridium
10	47	42.0	126	1	DOC_BPP1 Q06259 bacterioph
11	47	42.0	781	1	SP1_MOUSE O89090 mus musculu
12	47	42.0	785	1	SP1_MOUSE P08047 homo sapien
13	47	42.0	788	1	SP1_MOUSE O01714 rattus norv
14	46	41.1	231	1	DMH_AQUAE O67802 aquifex aeo
15	46	41.1	231	1	USF_AQUY P46209 aquifex pyr
16	46	41.1	245	1	TM4_METUA Q05723 methanococc
17	46	41.1	301	1	LECI_MOUSE P24721 mus musculu
18	45	40.2	587	1	MCCB_ARATH Q91d8 arabidopsis
19	44	39.3	1396	1	VITF_BPTS P13390 bacterioph
20	43	38.4	110	1	YC27_ARCVU O28041 archaeoglyc
21	43	38.4	299	1	Y002_CAEBL Q09529 caenoradid
22	43	38.4	467	1	ISBP_SCHPO P33984 schizosomy
23	41.5	37.1	401	1	HAT2_YEAST O02777 felis silve
24	41.5	37.1	472	1	CHIR_FELCA P21554 homo sapien
25	41.5	37.1	472	1	CHIR_HUMAN P47746 mus musculu
26	41.5	37.1	473	1	CHIR_MOUSE P20272 rattus norv
27	41.5	37.1	567	1	DY13_CHLRE P27766 chlamydom
28	41.5	37.1	579	1	COB2_BRARE O09375 brachydaric
29	41.5	37.1	336	1	VINT_BP186 P06723 bacterioph
30	41	36.6	412	1	DP42_BACSU P54560 bacillus su
31	41	36.6	412	1	DP41_BACSD O96c7 bacillus ha
32	41	36.6	413	1	DP41_BACSD O51831 porphyromon
33	41	36.6	433	1	MTRC_FORGI

34	41	36.6	473	1	CBIR_TARGR Q9pu17 taricha gra
35	41	36.6	513	1	Y073_TREPA O83112 treponema p
36	41	36.6	924	1	OSTA_PSEAE Q915u2 pseudomonas
37	40.5	36.2	464	1	VL2_HPV27 P36755 human papil
38	40.5	36.2	524	1	VL2_HPV27 P25487 human papil
39	40	35.7	205	1	3MGH_CUOPE O8xh9 clostridium
40	40	35.7	394	1	NUCC_STVY3 P27724 synochocyst
41	40	35.7	460	1	ALN_YEAST P23375 saccharomyc
42	40	35.7	467	1	RFAL_CRTFA Q23696 crithidia t
43	40	35.7	691	1	TOP1_BACSU P39814 bacillus su
44	40	35.7	719	1	WZC_SALTY O82566 salmonella
45	40	35.7	719	1	WZC_SALTY Q9f7b1 salmonella

ALIGNMENTS

```

RESULT 1
ID GTFB_STRMU STANDARD; PRT; 1476 AA.
AC P08987; O69381; O69384; O69387; O69390; O69396;
DT 01-NOV-1988 (Rel. 09, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Glucosyltransferase-I precursor (EC 2.4.1.5) (GTF-I) (Dextranucrase)
DE (Sucrose 6-glucosyltransferase).
GN GTFB OR SMU.1004.
OS Streptococcus mutans.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1309;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=GS-5;
RX MEDLINE=67308013; PubMed=3040685;
RA Shiroza T., Ueda S., Kuramitsu H.K.;
RT "Sequence analysis of the gtfb gene from Streptococcus mutans.";
RL J. Bacteriol. 169:4263-4270 (1987).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=MT4239 / Serotype C, MT4245 / Serotype E, MT4251 / Serotype F,
RX MT4467 / Serotype E, and MT8148 / Serotype C;
RA MEDLINE=98231643; PubMed=9570124;
RA Fujiwara T., Terao Y., Hoshino T., Kawabata S., Ooshima T., Sobue S.,
RA Kimura S., Hamada S.;
RT "Molecular analyses of glucosyltransferase genes among strains of
RX Streptococcus mutans.";
RL FEMS Microbiol. Lett. 161:331-336 (1998).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=UA159 / ATCC 700610 / Serotype C;
RX MEDLINE=22295063; PubMed=12397186;
RA Ajdic D., McMan W.M., McLaughlin R.E., Savic G., Chang J.,
RA Carson M.B., Primeaux C., Tian R., Kenton S., Jia H., Jia S., Qian Y.,
RA Li S., Zhu H., Najjar F., Lai H., White U., Roe B.A., Perreletti J.J.;
RT "Genome sequence of Streptococcus mutans UA159, a cariogenic dental
RX pathogen.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:14434-14439 (2002).
RN [4]
RP FUNCTION: PRODUCTION OF EXTRACELLULAR GLUCANS, THAT ARE THOUGHT
RX OF THEIR ABILITY TO ADHERE TO SMOOTH SURFACES AND MEDIANE THE
RX AGGREGATION OF BACTERIAL CELLS AND FOOD DEBRIS.
RX 1,3-LINKED GLUCOSE AND SOME 1,6 LINKAGES), GTF-S SYNTHESIZES
RX fructose + ((1,6)-alpha-D-glucosyl) (N+1).
RX AGGREGATION OF BACTERIAL CELLS AND FOOD DEBRIS.
RX 1,3-LINKED GLUCOSE AND SOME 1,6 LINKAGES), GTF-S SYNTHESIZES
RX WATER-SOLUBLE GLUCANS (ALPHA 1,6-GLUCOSE). GTF-SI SYNTHESIZES BOTH
RX FORMS OF GLUCANS.
RX - SIMILARITY: TO OTHER GLUCOSYLTRANSFERASES AND SOME TO A GLUCAN-
RX BINDING PROTEIN FROM S. MUTANS.
RX - SIMILARITY: Contains 10 cell wall binding repeats.

```

CC	-----
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration
CC	between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC	the European Bioinformatics Institute. There are no restrictions on its
CC	use by non-profit institutions as long as its content is in no way
CC	modified and this statement is not removed. Usage by and for commercial
CC	entities requires a license agreement (see http://www.isdb-sib.ch/announce/
CC	or send an email to license@isdb-sib.ch).
CC	-----
DR	EMBL; M17261; AAA895868.1; -
DR	EMBL; D88651; BAA26101.1; -
DR	EMBL; D88654; BAA26105.1; -
DR	EMBL; D88657; BAA26109.1; -
DR	EMBL; D88660; BAA26113.1; -
DR	EMBL; D89977; BAA26119.1; -
DR	EMBL; AE014940; AAN58705.1; -
DR	InterPro; IPR002479; CW binding.
DR	InterPro; IPR003318; Glyco_hydro. 70.
DR	Pfam; PF01473; CW binding_1; 1.
KW	Pfam; PF02324; Glyco_hydro_70; 1.
DW	Transferase; Glycosyltransferase; Signal; Repeat; Dental caries;
KM	Complete proteome.
FT	SIGNAL 1 34
FT	CHAIN 35 1476
FT	DOMAIN 35 1051
FT	DOMAIN 1097 1476
FT	REPEAT 1097 1130
FT	DOMAIN 1161 1470
FT	REPEAT 1161 1210
FT	REPEAT 1225 1275
FT	REPEAT 1290 1340
FT	REPEAT 1355 1405
FT	REPEAT 1420 1470
FT	REPEAT 62 62
FT	VARIANT 65 65
FT	VARIANT 68 68
FT	VARIANT 78 78
FT	VARIANT 86 86
FT	VARIANT 89 89
FT	VARIANT 168 168
FT	VARIANT 276 276
FT	VARIANT 399 399
FT	VARIANT 474 474
FT	VARIANT 512 512
FT	VARIANT 519 519
FT	VARIANT 701 701
FT	VARIANT 708 708
FT	VARIANT 938 938
FT	VARIANT 952 957
FT	VARIANT 963 964
FT	VARIANT 968 970
FT	VARIANT 1086 1086
FT	VARIANT 1158 1158
FT	VARIANT 1163 1163
FT	VARIANT 1168 1168
FT	VARIANT 1182 1182
FT	VARIANT 1234 1234
FT	VARIANT 1263 1263
FT	VARIANT 1263 1263
FT	VARIANT 1264 1264
FT	VARIANT 1272 1272
FT	VARIANT 1329 1329
FT	VARIANT 1394 1394
FT	VARIANT 1402 1402
CC	-----
CC	POTENTIAL.
CC	GLUCOSYLTRANSFERASE-T.
CC	CATALYTIC (APPROXIMATE).
CC	GLUCAN-BINDING (APPROXIMATE).
CC	A REPEAT.
CC	5 X TANDEM REPEATS.
CC	1.
CC	2.
CC	3.
CC	4.
CC	5.
CC	S -> T (IN STRAIN MT4239) .
CC	T -> I (IN STRAIN GS-5) .
CC	V -> A (IN STRAINS GS-5, MT4245, MT4251,
CC	MT4467 AND MT8148) .
CC	Q -> P (IN STRAIN MT4251) .
CC	I -> S (IN STRAINS GS-5, MT4245, MT4251,
CC	MT4467 AND MT8148) .
CC	S -> F (IN STRAIN MT4251) .
CC	X -> N (IN STRAIN MT4251) .
CC	S -> D (IN STRAINS GS-5, MT4467 AND
CC	MT8148) .
CC	N -> R (IN STRAIN MT4239) .
CC	I -> T (IN STRAIN MT4239) .
CC	K -> R (IN STRAIN MT8148) .
CC	F -> Y (IN STRAIN MT8148) .
CC	T -> I (IN STRAIN MT8148) .
CC	A -> V (IN STRAIN MT8148) .
CC	F -> L (IN STRAIN MT8148) .
CC	FGFGEV -> YGFPVA (IN STRAINS
CC	AND MT4467) .
CC	SV -> NT (IN STRAINS GS-5, MT4239 AND
CC	MT4467) .
CC	ADS -> VDG (IN STRAINS GS-5, MT4239 AND
CC	MT4467) .
CC	A -> T (IN STRAIN MT4239) .
CC	S -> N (IN STRAIN MT4239) .
CC	H -> Y (IN STRAIN MT4251) .
CC	E -> K (IN STRAIN MT8148) .
CC	Y -> C (IN STRAIN MT8148) .
CC	R -> P (IN STRAIN MT4239) .
CC	R -> E (IN STRAIN MT8148) .
CC	R -> H (IN STRAINS GS-5 AND MT4467) .
CC	Y -> H (IN STRAINS GS-5, MT4239, MT4467
CC	AND MT8148) .
CC	S -> G (IN STRAINS GS-5, MT4239, MT4467
CC	AND MT8148) .
CC	H -> Y (IN STRAINS GS-5 AND MT4467) .
CC	Y -> H (IN STRAINS GS-5, MT4239, MT4467
CC	AND MT8148) .
CC	S -> G (IN STRAINS GS-5, MT4239, MT4467

FT	VARIANT	1459	1459		AND MT814B).	
FT	CONFLICT	570	570		Y -> H (IN STRAIN MT4467).	
FT	CONFLICT	800	817		R -> A (IN REF. 1).	
FT	CONFLICT				ADDDVAAASTARSTDGK -> LHKFALRLAPHCQMA	
FT	CONFLICT				(IN REF. 1).	
FT	SEQUENCE	1310	1310		H -> L (IN REF. 1).	
SQ	SEQUENCE	1476	AA;	165846	MW; 9C6B95F731B4CBFC	CRC64;
Query Match				93.8%	Score 105;	DB 1; Length 1476;
Best Local Similarity				95.5%	Pred. No. 7,7e-09;	
Matches	21;	Conservative	0;	Mismatches	1;	Indels 0; Gaps 0;
Oy	1	TGARTINGOLLYPFRANGOVYKG	22			
Db	1301	TGARTINGOHLHYPRANGOVYKG	1322			
RESULT 2						
GTF2_STRDO						
ID	GTF2_STRDO	STANDARD;	PRT;	1592	AA.	
AC	P27470;					
DT	01-AUG-1992 (Rel. 23, Created)					
DT	01-AUG-1992 (Rel. 23, Last sequence update)					
DT	15-SEP-2003 (Rel. 42, Last annotation update)					
DT	GLUCOSYLTRANSFERASE-I precursor (EC 2.4.1.5) (GTF-I) (Dextranucrase)					
DE	(Sucrose 6-glucosyltransferase).					
DS	Streptococcus downei (Streptococcus sobrinus).					
OC	Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;					
CC	Streptococcus.					
OX	NCBI_TaxID=1317;					
RN	[1]					
RP	SEQUENCE FROM N.A.					
RC	STRAIN=6715 / Serotype G,					
RA	MEDLINE=91123227; PubMed=1704006;					
RA	Abo H., Matsumura T., Kodama T., Ohta H., Fukui K., Kato K.,					
RA	Kagawa H.;					
RT	"Peptide sequences for sucrose splitting and glucan binding within					
RT	Streptococcus sobrinus glucosyltransferase (water-insoluble glucan					
RL	synthetase)." J. Bacteriol. 173:989-996(1991).					
-I-	FUNCTION: PRODUCTION OF EXTRACELLULAR GLUCANS THAT ARE THOUGHT					
CC	TO PLAY A KEY ROLE IN THE DEVELOPMENT OF THE DENTAL PLAQUE BECAUSE					
CC	OF THEIR ABILITY TO ADHERE TO SMOOTH SURFACES AND MEDIANE THE					
CC	AGGREGATION OF BACTERIAL CELLS AND FOOD DEBRIS.					
-I-	CATALYTIC ACTIVITY: Sucrose + {(1,6)-alpha-D-glucosyl}(N) = D-					
CC	fructose + {(1,6)-alpha-D-glucosyl}(N+1).					
-I-	SUBCELLULAR LOCATION: Secreted.					
CC	-I- DISEASE: DENTAL CARIES.					
-I-	MICELLANEOUS: GTF-I SYNTHESIZES WATER-INSOLUBLE GLUCANS (ALPHA					
CC	1,3-LINKED GLUCOSE AND SOME 1,6 LINKAGES), GTF-S SYNTHESIZES					
CC	WATER-SOLUBLE GLUCANS (ALPHA 1,6-GLUCOSE). GTF-SI SYNTHESIZES BOTH					
CC	FORMS OF GLUCANS.					
-I-	SIMILARITY: TO OTHER GLUCOSYLTRANSFERASES AND SOME TO A GLUCAN-					
CC	BINDING PROTEIN FROM S.MUTANS.					
-I-	SIMILARITY: Contains 16 cell wall binding repeats.					
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration					
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -					
CC	the European Bioinformatics Institute. There are no restrictions on its use					
CC	by non-profit institutions as long as its content is in no way					
CC	modified and this statement is not removed. Usage by and for commercial					
CC	entities requires a license agreement (see http://www.isb-sdb.ch/announce/					
CC	or send an email to license@isb-sdb.ch).					
DR	EMBL; D90213; BAAl4241.1;					
DR	InterPro; IPR002479; CW_binding.					
DR	InterPro; IPR003318; Glyco_hydro_70.					
DR	Pfam; PF01473; CW_binding_1; 13.					
KM	Pfam; PF02324; Glyco_hydro_70; 1.					
FT	Transferase; Glucosyltransferase; Signal; Repeat; Dental caries.					
FT	SIGNAL	1	38			POTENTIAL.
FT	CHAIN	39	1592			GLUCOSYLTRANSFERASE-I.
FT	DOMAIN	39	1044			CATALYTIC (APPROXIMATE).

FT DOMAIN 1093 1592 GLUCAN-BINDING (APPROXIMATE).
 FT DOMAIN 1093 1592 6.5 X TANDEM REPEATS.
 FT REPEAT 1093 1142 1.
 FT REPEAT 1158 1207 2.
 FT REPEAT 1222 1272 3.
 FT REPEAT 1287 1337 4.
 FT REPEAT 1402 1451 5.
 FT REPEAT 1514 1563 6.
 FT REPEAT 1577 1592 7 (INCOMPLETE).
 SQ SEQUENCE 1592 AA; 176167 MM; BC0A6D079351ECP CRC64;

Query Match 77.7%; Score 87; DB 1; Length 1592;
 Best Local Similarity 77.3%; Pred. No. 6.6e-06;
 Matches 17; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 TGARTINGQULYFRANGQVKG 22
 Db 1298 TGAQTIRGQKLYFRANGQVKG 1319

RESULT 3
 GTF1_STRDO STANDARD; PRT; 1597 AA.
 ID GTF1_STRDO
 AC p11001:
 DT 01-JUL-1989 (Rel. 11, Created).
 DT 01-JUL-1989 (Rel. 11, Last sequence update).
 DT 15-SEP-2003 (Rel. 42, Last annotation update).
 DE Glucosyltransferase-1 precursor (EC 2.4.1.5) (GTF-1) (Dextranucrase)
 DE (sucrose 6-glucosyltransferase).
 GN GTF1.
 OS Streptococcus downei (Streptococcus sobrinus).
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 OX NCBI_Taxid=13117;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ME28;
 RX MEDLINE=87308014; PubMed=3040686;
 RA Ferretti J.J., Gilpin M.L., Russell R.R.B.;
 RT "Nucleotide sequence of a glucosyltransferase gene from Streptococcus
 sobrinus ME28".
 RL J. Bacteriol. 169:4271-4278 (1987).
 CC -1- FUNCTION: PRODUCTION OF EXTRACELLULAR GLUCANS, THAT ARE THOUGHT
 TO PLAY A KEY ROLE IN THE DEVELOPMENT OF THE DENTAL PLAQUE BECAUSE
 OF THEIR ABILITY TO ADHERE TO SMOOTH SURFACES AND MEDIANE THE
 AGGREGATION OF BACTERIAL CELLS AND FOOD DEBRIS.
 CC -1- CATALYTIC ACTIVITY: Sucrose + ((1,6)-alpha-D-glucosyl) (N) = D-
 fructose + ((1,6)-alpha-D-glucosyl) (N+1).
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- DISEASE: DENTAL CARIES.
 CC -1- MISCELLANEOUS: GTF-1 SYNTHESIZES WATER-INSOLUBLE GLUCANS (ALPHA
 1,3-LINKED GLUCOSE AND SOME 1,6 LINKAGES), GTF-S SYNTHESIZES
 WATER-SOLUBLE GLUCANS (ALPHA 1,6-GLUCOSE). GTF-SI SYNTHESIZES BOTH
 FORMS OF GLUCANS.
 CC -1- SIMILARITY: TO OTHER GLUCOSYLTRANSFERASES AND SOME TO A GLUCAN-
 BINDING PROTEIN FROM S. MUTANS.
 CC -1- SIMILARITY: Contains 19 cell wall binding repeats.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL, M17391, AAC63063.1; -
 DR InterPro; IPR002479; CW_binding.
 DR InterPro; IPR00318; Glyco_hydro_70.
 DR Pfam; PF01473; CW_binding_1; 16.
 DR Pfam; PF02324; Glyco_hydro_70; 1.
 KW Transferase; Glucosyltransferase; Signal; Repeat; Dental caries.
 FT SIGNAL 1 38 POTENTIAL.

FT CHAIN 39 1597 GLUCOSYLTRANSFERASE-1.
 FT DOMAIN 39 1050 CATALYTIC (APPROXIMATE).
 FT DOMAIN 1093 1597 GLUCAN-BINDING (APPROXIMATE).
 FT REPEAT 1093 1597 1.25 A, 2 B, AND 5 AC REPEATS.
 FT REPEAT 1099 1597 1.
 FT REPEAT 1163 1213 AC REPEAT.
 FT REPEAT 1227 1277 AC REPEAT.
 FT REPEAT 1292 1342 AC REPEAT.
 FT REPEAT 1352 1399 B REPEAT.
 FT REPEAT 1406 1455 AC REPEAT.
 FT REPEAT 1465 1512 B REPEAT.
 FT REPEAT 1519 1568 AC REPEAT.
 FT REPEAT 1582 1597 A REPEAT (INCOMPLETE).
 SQ SEQUENCE 1597 AA; 177080 MM; B9E86A2008687998 CRC64;

Query Match 77.7%; Score 87; DB 1; Length 1597;
 Best Local Similarity 77.3%; Pred. No. 6.6e-06;
 Matches 17; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 TGARTINGQULYFRANGQVKG 22
 Db 1303 TGAQTIRGQKLYFRANGQVKG 1324

RESULT 4
 GTF1_STRMU STANDARD; PRT; 1455 AA.
 ID GTF1_STRMU
 AC P33470; O69382; O69385; O69388; O69391; O69397; P05427;
 DT 01-NOV-1988 (Rel. 09, Created).
 DT 28-FEB-2003 (Rel. 41, Last sequence update).
 DT 15-SEP-2003 (Rel. 42, Last annotation update).
 DE Glucosyltransferase-51 precursor (EC 2.4.1.5) (GTF-SI)
 DE (Dextranucrase) (sucrose 6-glucosyltransferase).
 GN GTF1 OR SMU.1005.
 OS Streptococcus mutans.
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 OX NCBI_Taxid=1309;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=GS-5;
 RX MEDLINE=89137980; PubMed=2976010;
 RA Ueda S., Shiroza T., Kuramitsu H.K.;
 RT "Sequence analysis of the gtfC gene from Streptococcus mutans GS-5.";
 RL Gene 69:101-109 (1988).
 CC [2]
 CC -1- FUNCTION: PRODUCTION OF EXTRACELLULAR GLUCANS, THAT ARE THOUGHT
 TO PLAY A KEY ROLE IN THE DEVELOPMENT OF THE DENTAL PLAQUE BECAUSE
 OF THEIR ABILITY TO ADHERE TO SMOOTH SURFACES AND MEDIANE THE
 AGGREGATION OF BACTERIAL CELLS AND FOOD DEBRIS.
 CC -1- CATALYTIC ACTIVITY: Sucrose + ((1,6)-alpha-D-glucosyl) (N) = D-
 fructose + ((1,6)-alpha-D-glucosyl) (N+1).
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- DISEASE: DENTAL CARIES.
 CC -1- MISCELLANEOUS: GTF-1 SYNTHESIZES WATER-INSOLUBLE GLUCANS (ALPHA
 1,3-LINKED GLUCOSE AND SOME 1,6 LINKAGES), GTF-S SYNTHESIZES
 WATER-SOLUBLE GLUCANS (ALPHA 1,6-GLUCOSE). GTF-SI SYNTHESIZES BOTH
 FORMS OF GLUCANS.
 CC -1- SIMILARITY: TO OTHER GLUCOSYLTRANSFERASES AND SOME TO A GLUCAN-
 BINDING PROTEIN FROM S. MUTANS.
 CC -1- SIMILARITY: Contains 19 cell wall binding repeats.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL, M17391, AAC63063.1; -
 DR InterPro; IPR002479; CW_binding.
 DR InterPro; IPR00318; Glyco_hydro_70.
 DR Pfam; PF01473; CW_binding_1; 16.
 DR Pfam; PF02324; Glyco_hydro_70; 1.
 KW Transferase; Glucosyltransferase; Signal; Repeat; Dental caries.
 FT SIGNAL 1 38 POTENTIAL.

	CC	-I- CATALYTIC ACTIVITY: Sucrose + {(1,6)-alpha-D-glucosyl} (N) = D-factose + {(1,6)-alpha-D-glucosyl} (N+1).
	CC	-I- SUBCELLULAR LOCATION: Secreted.
	CC	-I- DISEASE: DENTAL CARIES.
	CC	-I- MISCELLANEOUS: GTF-I SYNTHESIZES WATER-IN-SOLUBLE GLUCANS (ALPHA 1,3-LINKED GLUCOSE AND SOME 1,6 LINKAGES), GTF-S SYNTHESIZES WATER-SOLUBLE GLUCANS (ALPHA 1,6-GLUCOSE). GTF-SI SYNTHESIZES BOTH FORMS OF GLUCANS.
	CC	-I- SIMILARITY: TO OTHER GLUCOSYLTRANSFERASES AND SOME TO A GLUCAN-BINDING PROTEIN FROM S. MUTANS.
	CC	-I- SIMILARITY: Contains 5 cell wall binding repeats.
	CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
	CC	-----
DR	EMBL	M22054; AAA89592.1; .
DR	EMBL	D88652; BAA26102.1; .
DR	EMBL	D88655; BAA26106.1; .
DR	EMBL	D88658; BAA26110.1; .
DR	EMBL	D88651; BAA26114.1; .
DR	EMBL	D89978; BAA26120.1; .
DR	EMBL	A601940; AAN58706.1; .
DR	EMBL	MF7361; AAA88589.1; .
DR	InterPro	IPIR003479; CW binding.
DR	InterPro	IPIR003348; Glyco_hydro_70.
DR	Pfam	PF01473; CW binding 1; 1.
KW	Pfam	PF02324; Glyco_hydro_70; 1.
	Transferrase	Glycosyltransferase; Signal; Repeat; Dental caries;
XW	Complete proteome	.
FT	SIGNAL	1 34
FT	CHAIN	35 1455
FT	DOMAIN	35 1050
FT	DOMAIN	1126 1455
FT	REPEAT	1126 1455
FT	REPEAT	1126 1159
FT	REPEAT	1169 1200
FT	REPEAT	1227 1238
FT	REPEAT	1253 1303
FT	REPEAT	1318 1330
FT	VARIANT	21 21
FT	VARIANT	81 81
FT	VARIANT	106 106
FT	VARIANT	116 116
FT	VARIANT	126 126
FT	VARIANT	150 151
FT	VARIANT	256 256
FT	VARIANT	425 425
FT	VARIANT	519 519
FT	VARIANT	538 538
FT	VARIANT	545 545
FT	VARIANT	597 597
FT	VARIANT	600 600
FT	VARIANT	601 601
FT	VARIANT	614 614
FT	VARIANT	727 727
FT	VARIANT	734 734
FT	VARIANT	964 964
FT	VARIANT	1113 1113
FT	VARIANT	1118 1118
FT	VARIANT	1204 1204
FT	VARIANT	1208 1208
	GLUCOSYLTRANSFERASE-ST,	CATALYTIC (APPROXIMATE).
	GLUCOGEN-BINDING (APPROXIMATE).	2,4-A, 1-C AND 1 AC REPEATS.
	A REPEAT.	A REPEAT.
	C REPEAT.	C REPEAT.
	AC REPEAT.	AC REPEAT.
	(INCOMPLETE).	A REPEAT (INCOMPLETE).
	V-> I (IN STRAIN GS-5).	V-> L (IN STRAIN MT4239).
	P-> L (IN STRAIN GS-5).	D-> V (IN STRAIN GS-5).
	S-> A (IN STRAINS GS-5 AND MT4467).	A-> T (IN STRAIN GS-5).
	SR-> PK (IN STRAINS GS-5, MT4239 AND MT4467).	R-> K (IN STRAINS MT4245, MT4251, MT4467 AND MT8148).
	R-> V (IN STRAINS GS-5 AND MT4467).	R-> T (IN STRAIN GS-5).
	A->> T (IN STRAIN GS-5).	T-> I (IN STRAIN GS-5).
	M->> I (IN STRAIN GS-5).	L->> F (IN STRAIN MT8148).
	X->> K (IN STRAINS MT4245 AND MT4251).	N->> Y (IN STRAIN MT4239).
	N->> T (IN STRAIN MT4239).	I->> V (IN STRAINS GS-5, MT4239, MT4467 AND MT8148).
	V->> I (IN STRAIN MT8148).	

```
CC FT VARIANT      1292   1294                                     DGH -> NOY (IN STRAINS GS-5, MT4467 AND  
FT FT VARIANT     1305   1369                                MISSING (IN STRAIN MT4245).  
FT FT VARIANT     1326   1326                                T -> V (IN STRAIN MT8148) .  
FT FT VARIANT     1331   1331                                I -> A (IN STRAINS GS-5 , MT4239 , MT4467  
                                 AND MT8148 ) .  
FT FT VARIANT     1377   1377                                R -> K (IN STRAIN MT8148) .  
FT FT VARIANT     1398   1398                                V -> I (IN STRAIN MT8148) .  
FT FT VARIANT     1424   1424                                D -> N (IN STRAIN MT4239) .  
FT FT VARIANT     1439   1439                                V -> I (IN STRAINS MT4239 AND MT8148) .  
FT FT VARIANT     1444   1444                                S -> P (IN STRAIN MT8148) .  
FT FT CONFLICT    1337   1455                                ORLYPSNGVOAGALKTEKGRIKXYDDPNSGENVNRKYR  
                                  TSQGMWYPENDDCAYALIGMVFRRGPRYPDENGVIYASHD  
                                  QRNHWIDYRDSDRGSSSAVRFRHRSGFDFNFPF ->  
                                  HASILSLMWRLRESSLQSXKVYSNMIIILEPKAFIVM  
                                  (IN REF. 1) .
```

```
CC SQ SEQUENCE     1455 AA; 162965 MW; 3CB455A9A4FECD6 CRC64;
```

```
CC Query Match           76.8%; Score 86; DB 1; Length 1455;  
CC Best Local Similarity 68.2%; Pred. No. 8.7e-06;  
CC Matches 15; Conservative 3; Mismatches 4; Indels 0; Gaps 0;
```

```
Cy          1 TGARTINGQLYLIFRANGYOVKG 22  
Db          1329 TGTRVVVGQRLLTFFKSNGVQAKG 1350
```

```
--|---|||----||--|||||--|  
-----|:::~::~|||---
```

```
GFES_STRDO RESULT 5  
ID _GFES_STRDO STANDARD; PRS; 1365 AA.  
AC P29336;  
DT 01-DEC-1992 (Rel. 24, Created)  
DT 01-DEC-1992 (Rel. 24, Last sequence update)  
DT 28-FEB-2003 (Rel. 4), Last annotation update)  
DE Glucosyltransferase-S precursor (EC 2.4.1.5) (GTF-S) (Dextranucrase)  
DE (Sucrose 6-glucosyltransferase) .  
GN GTF.S.
```

```
OS Streptococcus downei (Streptococcus sobrius).  
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;  
OX Streptococcus;  
ON NCBI_TaxId=1317;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=MFBZ8;  
RA MEDLINE=90316665; PubMed=2142479;  
RX Gilmore K.S., Russell R.R., Ferretti J.J.;  
RT "Analysis of the Streptococcus downei gtfS gene, which specifies a  
RT glucosyltransferase that synthesizes soluble glucans." ;  
RL Infect Immun. 58:2452-2458 (1990).
```

```
-I- FUNCTION: PRODUCTION OF EXTRACELLULAR GLUCANS, THAT ARE THOUGHT  
TO PLAY A KEY ROLE IN THE DEVELOPMENT OF THE DENTAL PLAQUE BECAUSE  
OF THEIR ABILITY TO ADHERE TO SMOOTH SURFACES AND Mediate the  
ACCUMULATION OF BACTERIAL CELLS AND FOOD DEBRIS.  
-I- CATALYTIC ACTIVITY: Sucrose + [(1,6)-alpha-D-glucoeryl] (N)+1 = D-  
fructose + [(1,6)-alpha-D-glucoeryl] (N+1).  
-I- ENZYME REGULATION: GLUCAN SYNTHESIS BY GTF-S IS INDEPENDENT OF  
PRIMER GLUCAN UNLIKE GTF-I.  
-I- DISASSOCIATES DENTAL CARIES.  
-I- MISCELLANEOUS: GTF-S SYNTHESIZES WATER-SOLUBLE GLUCANS (ALPHA  
1,6-GULOSE).  
-I- SIMILARITY: TO OTHER GLUCOSYLTRANSFERASES AND SOME TO A GLUTAN-  
BINDING PROTEIN FROM S. MUTANS.  
-I- SIMILARITY: Contains 10 cell wall binding repeats.
```

```
=====  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation at  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way dis-  
CC seminated and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/  
or send an email to license@isb-sib.ch).  
EMBL; M30943; AAA6898.1; -. 
```

```

DR InterPro; IPR002479; CW binding.
DR InterPro; IPR003318; Glyco_hydro_70.
DR Pfam; PF01473; CW binding 1; 8.
DR Pfam; PF02324; Glyco_hydro_70; 1.
KW Transferase; Glycosyltransferase; Signal; Repeat; Dental caries.
FT SIGNAL 1 36
FT CHAIN 37 1365
FT REPEAT 157 177
FT REPEAT 178 197
FT DOMAIN 198 1061
FT REPEAT 1062 1082
FT REPEAT 1083 1102
FT REPEAT 1150 1169
FT REPEAT 1170 1190
FT REPEAT 1225 1243
FT REPEAT 1289 1308
FT REPEAT 1309 1328
FT REPEAT 1331 1352
SQ SEQUENCE 1365 AA; 151590 MW; 167296B5A2EBC476 CRC64;

Query Match 67.0%; Score 75; DB 1; Length 1365;
Best Local Similarity 63.6%; Pred. No. 0.00048;
Matches 14; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

QY 1 TGARTINGOLLYPRANGQVKG 22
DB 1235 TGEOTIDGKVFQDNGVQVKG 1256

RESULT 6
GTFD_STRMU STANDARD; PRT; 1462 AA.
AC P4331; O69383; O69386; O69389; O69392; O69398;
DT 01-FEB-1996 (Rel. 33, Created)
DT 26-FEB-2003 (Rel. 41, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Glucosyltransferase-S precursor (EC 2.4.1.5) (GTF-S) (Dextranucrase)
DE (Sucrose 6-glucosyltransferase).
GN GTFD OR SMU. 910.
OS Streptococcus mutans.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1309;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=GS-5;
RX MEDLINE=91100958; PubMed=2148600;
RA Honda O., Kato C., Kuramitsu H.K.;
RT "Nucleotide sequence of the Streptococcus mutans gtfD gene encoding
   the glucosyltransferase-S enzyme."
RL J. Gen. Microbiol. 136:2099-2105(1990).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=MT4239 / Serotype C, MT4245 / Serotype E, MT4251 / Serotype F,
RX MT4467 / Serotype E, and MT8148 / Serotype C;
RA MEDLINE=98231643; PubMed=9570124;
RA Fujiwara T., Teruo Y., Hoshino T., Kawabata S., Ooshima T., Sobue S.,
   Kimura S., Hamada S.;
RT "Molecular analyses of glucosyltransferase genes among strains of
   Streptococcus mutans."
RL FEMS Microbiol. Lett. 161:331-336(1998).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=UA159 / ATCC 700610 / Serotype C;
RX MEDLINE=22295063; PubMed=12397186;
RA Ajdidi D., McShan W.M., McLaughlin R.E., Savic G., Chang J.,
   Carson M.B., Prineaux C., Tian R., Kenton S., Jia H., Lin S., Qian Y.,
   Li S., Zhu H., Nejar F., Lai H., White J., Roe B.A., Ferretti J.J.;
RT "Genome sequence of Streptococcus mutans UA159, a cariogenic dental
   pathogen."
RL Proc. Natl. Acad. Sci. U.S.A. 99:14434-14439(2002).
CC -!- FUNCTION: PRODUCTION OF EXTRACELLULAR GLUCANS, THAT ARE THOUGHT
   TO PLAY A KEY ROLE IN THE DEVELOPMENT OF THE DENTAL PLAQUE BECAUSE

```

```

CC OF THEIR ABILITY TO ADHERE TO SMOOTH SURFACES AND MEDIATE THE
CC AGGREGATION OF BACTERIAL CELLS AND FOOD DEBRIS.
CC -!- CATALYTIC ACTIVITY: Sucrose + ((1,6)-alpha-D-glucosyl) (N) = D-
CC fructose + ((1,6)-alpha-D-glucosyl) (N+1).
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- DISEASE: DENTAL CARIES.
CC -!- MISCELLANEOUS: GTF-S SYNTHESIZES WATER-INSOLUBLE GLUCANS (ALPHA
CC 1,3-LINKED GLUCOSE AND SOME 1,6 LINKAGES). GTF-S SYNTHESIZES
CC WATER-SOLUBLE GLUCANS (ALPHA 1,6-GLUCOSE). GTF-SI SYNTHESIZES BOTH
CC FORMS OF GLUCANS.
CC -!- SIMILARITY: TO OTHER GLUCOSYLTRANSFERASES AND SOME TO A GLUCAN-
CC BINDING PROTEIN FROM S. MUTANS.
CC -!- SIMILARITY: Contains 6 cell wall binding repeats.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M29296; AAA26895.1; -
CC EMBL; D88653; BAA26103.1; -
CC EMBL; D88656; BAA26107.1; -
CC EMBL; D88658; BAA26111.1; -
CC EMBL; D88662; BAA26115.1; -
CC EMBL; D89979; BAA26121.1; -
CC EMBL; AE014932; AAN58619.1; -
CC InterPro; IPR002479; CW binding.
CC InterPro; IPR003318; Glyco_hydro_70.
CC Pfam; PF02324; Glyco_hydro_70; 1_70.
CC Transferase; Glycosyltransferase; Signal; Repeat; Dental caries;
CC KW Complete proteome.
CC FT SIGNAL 1 1462
CC FT CHAIN 2 1462
CC FT DOMAIN 1232 1423
CC FT REPEAT 1232 1295
CC FT REPEAT 1296 1355
CC FT REPEAT 1360 1423
CC FT REPEAT 10 10
CC FT VARIANT 19 19
CC FT VARIANT 58 58
CC FT VARIANT 68 68
CC FT VARIANT 81 81
CC FT VARIANT 113 113
CC FT VARIANT 122 122
CC FT VARIANT 132 132
CC FT VARIANT 135 135
CC FT VARIANT 137 137
CC FT VARIANT 202 202
CC FT VARIANT 255 255
CC FT VARIANT 275 275
CC FT VARIANT 288 288
CC FT VARIANT 301 301
CC FT VARIANT 313 313
CC FT VARIANT 317 317
CC FT VARIANT 328 328
CC FT VARIANT 350 350
CC FT VARIANT 628 633
CC FT VARIANT 688 688
CC FT VARIANT 726 732
CC FT VARIANT 762 762
CC FT VARIANT 762 762

```

```

FT  VARIANT  964  964  D -> Y (IN STRAIN MT4251).
FT  VARIANT  1019 1019  E -> K (IN STRAINS MT4245 AND MT4251).
FT  VARIANT  1059 1060  LG -> IR (IN STRAIN MT4251).
FT  VARIANT  1060 1060  G -> R (IN STRAIN MT4245).
FT  VARIANT  1080 1080  G -> R (IN STRAIN MT4239).
FT  VARIANT  1142 1142  H -> Q (IN STRAIN GS-5).
FT  VARIANT  1198 1198  S -> N (IN STRAIN MT4239).
FT  VARIANT  1220 1220  Y -> C (IN STRAINS MT4251 AND MT4467).
FT  VARIANT  1280 1280  Y -> L (IN STRAIN MT4467).
FT  VARIANT  1282 1282  Q -> P (IN STRAIN MT4245).
FT  VARIANT  1290 1290  K -> T (IN STRAIN MT4245).
FT  VARIANT  1311 1311  N -> D (IN STRAIN MT4245).
FT  VARIANT  1403 1403  G -> D (IN STRAINS GS-5 AND MT4467).
FT  VARIANT  1425 1425  G -> R (IN STRAIN GS-5).
FT  VARIANT  1449 1449  R -> K (IN STRAIN MT4467).
FT  CONFLICT 1428 1462  R (IN REF. 1).
SQ  SEQUENCE 1462 AA; 163387 MW; CE4A279C4D708645 CRC64;

Query Match 59.8%; Score 67; DB 1; Length 1462;
Best Local Similarity 59.1%; Pred. No. 0.01;
Matches 13; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

Oy 1 TGARTINGQLLYFRANGVQVKG 22
Db 1332 TGSQTIAKTKLYFASDQVQVKG 1353

RESULT 7
Y19 NPVOP STANDARD; PRT; 529 AA.
ID Y19 NPVOP STANDARD; PRT; 529 AA.
AC 010358;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DE Hypothetical 59.0 kDa protein precursor (ORF119).
OS Orygia pseudotsugata multicapsid polyhedrosis virus (OpMNPV).
OC Viruses; dsDNA viruses, no RNA stage; Baculoviridae;
OC Nucleopolyhedrovirus.
OX NCBI_TaxID=164623;
[1]
RP MEDLINE=97271300; PubMed=9126251;
RA Ahrens C.H., Russell R.R., Funk C.J., Evans J., Harwood S.,
RA Rohkman G.F.;
RT "The sequence of the Orygia pseudotsugata multinucleocapsid nuclear
RT polyhedrosis virus genome."
RL Virology 229:381-393(1997).
CC -1- SIMILARITY: TO CORRESPONDING ORF IN ACNMPV.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U75930; AAC5918.1; -.
DR Pfam; PF05092; DUF686; 1.
FT SIGNAL 1 20
FT CHAIN 21 529 HYPOTHETICAL PROTEIN ORF119.
SQ SEQUENCE 529 AA; 58983 MW; E85F81DD9219BCB CRC64;

Query Match 44.2%; Score 49.5; DB 1; Length 529;
Best Local Similarity 50.0%; Pred. No. 2.4;
Matches 11; Conservative 5; Mismatches 5; Indels 1; Gaps 1;

Oy 1 TGARTINGQLLYFRANGVQVKG 22
Db 260 SGRRT-SGRLEFYHADGVEVSG 280

```

```

RESULT 8
ID TOXA CLODI STANDARD; PRT; 2710 AA.
ID TOXA CLODI STANDARD; PRT; 2710 AA.
AC P16154;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Toxin A.
GN TOXA OR TCDA.
OS Clostridium difficile.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1496;
[1]
RP SEQUENCE FROM N.A.
RP STRAIN=VPI 10463;
RC MEDLINE=90221894; PubMed=2109310;
RX Sauerborn M., von Eichel-Streiber C.;
RA "Nucleotide sequence of Clostridium difficile toxin A."
RT Nucleic Acids Res. 18:1629-1630(1990).
[2]
RP SEQUENCE FROM N.A.
RP STRAIN=VPI 10463;
RC MEDLINE=90129305; PubMed=2105276;
RX Dore C.H., Wang S.Z., Price S.B., Phelps C.J., Lysterly D.M.,
RA Wilkins T.W., Johnson J.L.;
RT "Molecular characterization of the Clostridium difficile toxin A
RT gene."
RL Infect. Immun. 58:480-488(1990).
[3]
RP SEQUENCE FROM N.A.
RP STRAIN=VPI 10463;
RA von Eichel-Streiber C.;
RT Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.
CC -1- DOMAIN: THE C-TERMINAL PART OF TOXIN A CONSISTS OF A 83 AA
CC REPETITIVE STRUCTURE. THIS PART OF TOXIN A IS COMPOSED OF FIVE
CC DIFFERENT OLIGOPEPTIDES.
CC -1- DISEASE: CLOSTRIDIUM DIFFICILE PRODUCES TWO TOXINS, AN
CC ENTEROTOXIN CALLED A AND CYTOTOXIN B. ONLY AFTER THE ENTERAL
CC DELIVERY OF THE ENTEROTOXIN A MAY THE CHARACTERISTIC DISEASE
CC CALLED PSEUDOMEMBRANOUS COLITIS BE INDUCED.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X51797; CAA36094.1; -.
DR EMBL; M30307; AAA23283.1; -.
DR EMBL; X92982; CAA63564.1; -.
DR InterPro; IPR002479; CW binding.
DR Pfam; PR01473; CW binding 1; 28.
DR Pfam; PR04488; Gly_transf_sug; 1.
KW Toxin; Enterotoxin.
SQ SEQUENCE 2710 AA; 308052 MW; 0A6E52CE84C14421 CRC64;

Query Match 44.2%; Score 49.5; DB 1; Length 2710;
Best Local Similarity 52.2%; Pred. No. 13;
Matches 12; Conservative 2; Mismatches 8; Indels 1; Gaps 1;

Oy 1 TGARTIDGKKYFNTTAAVVG 2462
Db 2460 TGARTIDGKKYFNTTAAVVG 2462

RESULT 9
BIOD_CLOAB STANDARD; PRT; 240 AA.
ID BIOD_CLOAB STANDARD; PRT; 240 AA.
AC Q97UC5;

```



```

CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=1;
CC IsoId=089090-1; Sequence=Displayed;
CC Name=2; Synonyms=Spl-1;
CC IsoId=089090-2; Sequence=VSP 007376;
CC -1- PTM: O-GLYCOSYLATED; CONTAINS N-ACETYLGLUCOSAMINE SIDE CHAINS
CC (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE SPL FAMILY OF C2H2-TYPE ZINC-FINGER
CC PROTEINS.
CC -1- SIMILARITY: Contains 3 C2H2-type zinc fingers.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF062566; AAC16484.1; -
CC EMBL; S79832; AAB35321.1; -
CC EMBL; X60136; CAA42721.1; -
CC HSSP; P08047; 1SP1.
CC MGD; MGI:98372; Spl.
CC GO; GO:0005634; C:nucleus; IDA.
CC GO; GO:0003700; F:transcription factor; IMP.
CC GO; GO:000355; P:regulation of transcription, DNA-dependent; IMP.
CC InterPro: IPR007087; Znf_C2H2.
CC Pfam: PF00096; zf-C2H2; 3.
CC ProDom: PD000003; Znf_C2H2; 2.
CC SMART; SM00355; Znf_C2H2; 3.
CC PROSITE; PS00028; ZINC_FINGER_C2H2_1; 3.
CC PROSITE; PS00157; ZINC_FINGER_C2H2_2; 3.
CC Transcription regulation; Activator; Zinc-finger; Metal-binding;
CC DNA-binding; Nuclear protein; Repeat; Glycoprotein;
CC KW Alternative splicing.
CC FT ZN_FING 624 648 C2H2-TYPE 1.
CC FT ZN_FING 654 678 C2H2-TYPE 2.
CC FT ZN_FING 684 706 C2H2-TYPE 3.
CC FT VARSP LIC 57 370 Missing (in isoform 2).
CC FT VARSP LIC 57 370 /FridaVSP 007376.
CC FT CONFLICT 459 459 V -> G (in Ref. 2).
CC SQ SEQUENCE 781 AA; 80486 MW; 14CD12B8C58CF21 CRC64;
CC Query Match 42.0%; Score 47; DB 1; Length 781;
CC Best Local Similarity 50.0%; Pred. No. 8.9;
CC Matches 9; Conservative 4; Mismatches 5; Indels 0; Gaps 0;
QY 4 RTINGQULIFRANGVQVK 21
Db 179 QTVDGQQLQPAATGAQVQ 196

```

```

RN [2]
RP SEQUENCE OF 1-558 FROM N.A.
RX MEDLINE=20545561; PubMed=10973950;
RA Takahara T., Kanazu S., Yanagisawa S., Akenuma H.,
RT "Heterogeneous Spl mRNAs in human Hep2 cells include a product of
RT homotypic trans-splicing."
RL J. Biol. Chem. 273:38067-38072(2000).
RN [3]
RP SEQUENCE OF 90-785 FROM N.A., AND SEQUENCE OF 359-375 AND 670-675.
RX MEDLINE=8080466; PubMed=3319186;
RA Kadonaga J.T., Garner K.R., Masiarz F.R., Tjian R.,
RT "Isolation of cDNA encoding transcription factor Spl and functional
RT analysis of the DNA binding domain."
RL Cell 51:1079-1090(1987).
RN [4]
RP SEQUENCE OF 1-109 FROM N.A.
RA Nicolas M., Noe V., Ciudad C.J.,
RT "Expression of transcription factor Spl mRNA in mammalian cells."
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE OF 1-98 FROM N.A.
RA Handebug K., Huebner A.,
RT "Sequencing of the 5' end of human transcription factor Spl mRNA."
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
RN [6]
RP O-GLYCOSYLATION.
RX MEDLINE=8903041; PubMed=3139301;
RA Jackson S.P., Tjian R.,
RT "O-glycosylation of eukaryotic transcription factors: implications
RT for mechanisms of transcriptional regulation."
RL Cell 55:125-133(1988).
RN [7]
RP STRUCTURE BY NMR OF 654-684 AND 684-712.
RX MEDLINE=97218212; PubMed=9065444;
RA Narayan V.A., Kiriacki R.W., Caradonna J.P.,
RT "Structures of zinc finger domains from transcription factor Spl.
RT Insights into sequence-specific protein-DNA recognition."
RL J. Biol. Chem. 272:7801-7809(1997).
RN [8]
RP IDENTIFICATION OF SEROTONIN 1A RECEPTOR PROMOTER BINDING SITES.
RX MEDLINE=96224025; PubMed=8626793;
RA Parks C.L., Shenk T.,
RT "The serotonin 1a receptor gene contains a TATA-less promoter that
RT responds to MAZ and Spl."
RL J. Biol. Chem. 271:4417-4430(1996).
RN [9]
RP FUNCTION: BINDS TO GC BOX PROMOTERS ELEMENTS AND SELECTIVELY
CC ACTIVATES MRNA SYNTHESIS FROM GENES THAT CONTAIN FUNCTIONAL
CC RECOGNITION SITES. CAN INTERACT WITH G/C-RICH MOTIFS FROM
CC SEROTONIN RECEPTOR PROMOTER.
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- PTM: O-GLYCOSYLATED; CONTAINS N-ACETYLGLUCOSAMINE SIDE CHAINS.
CC -1- SIMILARITY: BELONGS TO THE SPL FAMILY OF C2H2-TYPE ZINC-FINGER
CC PROTEINS.
CC -1- SIMILARITY: Contains 3 C2H2-type zinc fingers.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF252284; AAF67726.1; -
CC EMBL; AB039286; BAB13476.1; -
CC EMBL; J03133; AAA61154.1; -
CC EMBL; AF255682; AAF78781.1; -
CC EMBL; AJ272134; CAB75345.1; -
CC PIR; A29635; A29635.
CC PDB; 1SP1; 21-APR-97.
CC PDB; 1SP2; 21-APR-97.
CC TRANSFAC; T00759; -
CC GlycoSiteDB; P08047; -

```

DR Genew; HGNC:11205; SPL.
 DR MTM; 189905; -
 DR GO; GO:0005634; C:nucleus; NAS.
 DR GO; GO:0003702; F:RNA polymerase II transcription factor acti. .; NAS.
 DR GO; GO:0016563; F:transcriptional activator activity; NAS.
 DR GO; GO:0006355; P:regulation of transcription; DNA-dependent; NAS.
 DR InterPro; IPR007087; Znf C2H2.
 DR Pfam; PF00096; zf-C2H2; 3.
 DR Prodom; PD000003; Znf C2H2; 2.
 DR SMART; SM00355; Znf C2H2; 3.
 DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 3.
 DR PROSITE; PS00157; ZINC_FINGER_C2H2_2; 3.
 DR Transcription regulation; Activator; Zinc-finger; Metal-binding;
 KW DNA-binding; Nuclear protein; Repeat; Glycoprotein; 3D-structure.
 FT ZN_FING 626 650
 FT ZN_FING 656 680 C2H2-TYPE 2.
 FT ZN_FING 686 708 C2H2-TYPE 3.
 FT CONFLICT 366 366 D -> G (IN REF. 3; AA SEQUENCE).
 FT CONFLICT 670 670 S -> F (IN REF. 3; AA SEQUENCE).
 FT STRAND 657 657
 FT STRAND 661 662
 FT STRAND 666 666
 FT HELIX 670 677
 FT TURN 678 680
 SQ SEQUENCE 785 AA; 80693 MW; 43893DBF6518B9EA CRC64;
 Query Match 42.0%; Score 47; DB 1; Length 785;
 Best Local Similarity 50.0%; Pred. No. 8.9;
 Matches 9; Conservative 4; Mismatches 5; Indels 0; Gaps 0;
 QY 4 RTINGQLLYFRANGVQYK 21
 Db 180 QTVNGQQLQFPAATGAQVQ 197
 RESULT 13
 SPL_RAT STANDARD; PRT; 788 AA.
 AC Q01714;
 DT 01-APR-1993 (Rel. 25, Created)
 DT 01-APR-1993 (Rel. 25, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Transcription factor Spl.
 GN Spl.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Liver;
 RX MEDLINE=93010958; PubMed=1356762; Kikuchi Y., Sasano K.,
 RA Imataka H., Sogawa K., Yasumoto K., Fujii-Kuriyama Y.,
 RA Kobayashi A., Hayami M., Fujii-Kuriyama Y.,
 RT "Two regulatory proteins that bind to the basic transcription element
 (BME), a GC box sequence in the promoter region of the rat P-4501A1
 gene.";
 RL EMBL J. 11:3663-3671 (1992).
 CC -1- FUNCTION: BINDS TO GC BOX PROMOTERS ELEMENTS AND SELECTIVELY
 CC ACTIVATES MRNA SYNTHESIS FROM GENES THAT CONTAIN FUNCTIONAL
 CC RECOGNITION SITES.
 CC -1- SUBCELLULAR LOCATION: Nuclear.
 CC -1- PTM: O-GLYCOSYLATED; CONTAINS N-ACETYLGLUCOSAMINE SIDE CHAINS
 CC (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE SPL FAMILY OF C2H2-TYPE ZINC-FINGER
 CC PROTEINS.
 CC -1- SIMILARITY: Contains 3 C2H2-type zinc fingers.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; D12768; BAA02235.1; -
 DR PIR; J050747; J050747.
 DR HSSP; P08047; ISPL.
 DR TRANSFAC; T00754; -
 DR InterPro; IPR007087; Znf C2H2.
 DR Pfam; PF00096; zf-C2H2; 3.
 DR Prodom; PD000003; Znf C2H2; 2.
 DR SMART; SM00355; Znf C2H2; 3.
 DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 3.
 DR PROSITE; PS00157; ZINC_FINGER_C2H2_2; 3.
 DR Transcription regulation; Activator; Zinc-finger; Metal-binding;
 KW DNA-binding; Nuclear protein; Repeat; Glycoprotein.
 FT ZN_FING 629 653
 FT ZN_FING 659 683 C2H2-TYPE 2.
 FT ZN_FING 689 711 C2H2-TYPE 3.
 SQ SEQUENCE 788 AA; 81015 MW; AA2B0CAB81AAB80C CRC64;
 Query Match 42.0%; Score 47; DB 1; Length 788;
 Best Local Similarity 50.0%; Pred. No. 9;
 Matches 9; Conservative 4; Mismatches 5; Indels 0; Gaps 0;
 QY 4 RTINGQLLYFRANGVQYK 21
 Db 183 QTVNGQQLQFPAATGAQVQ 200
 RESULT 14
 DLH_AQUAE STANDARD; PRT; 231 AA.
 AC Q67802;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Putative carboxymethyl-enebutenolidease (Ec 3.1.1.45) (Dienelactone
 DE hydrolase) (DLH).
 GN AQ 1997.
 OS Aquifex aeolicus.
 OC Bacteria; Aquificae; Aquificales; Aquificaceae; Aquifex.
 OX NCBI_TaxID=63363;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=VPS;
 RX MEDLINE=9819666; PubMed=9537320;
 RA Deckert G., Warren P.V., Gaasterland T., Young W.G., Lenox A.L.,
 RA Graham D.B., Overbeek R., Snead M.A., Keller M., Ausley M., Huber R.,
 RA Feldman R.A., Short J.M., Olson G.J., Swanson R.V.;
 RT "The complete genome of the hyperthermophilic bacterium Aquifex
 RT aeolicus.";
 RL Nature 392:353-358 (1998).
 CC -1- CATALYTIC ACTIVITY: 4-carboxymethyl-enebut-2-en-4-olide + H(2)O = 4
 CC oxohe-2-enebutate.
 CC -1- SIMILARITY: BELONGS TO THE DIENELACTONE HYDROLASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AE000767; AAC07773.1; -
 DR PIR; F70471; F70471.
 DR InterPro; IPR002925; DLH.
 DR InterPro; IPR00379; Ser_estr_ site.
 DR Pfam; PF01738; DLH; 1.
 KW Hypothetical protein; Hydrolase; Complete proteome.
 FT ACT_SITE 118 118 BY SIMILARITY.
 FT ACT_SITE 167 167 BY SIMILARITY.
 FT ACT_SITE 199 199 BY SIMILARITY.

SQ SEQUENCE 231 AA; 26357 MW; 8268DA32A6980751 CRC64;

Query Match 41.1%; Score 46; DB 1; Length 231;
 Best Local Similarity 53.3%; Pred. No. 3.7;
 Matches 8; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

OY 8 GOLLVFRANGVQYKG 22
 ||::| ||::|
 Db 2 GQVVEFEKNGVQVKG 16

RESULT 15

USF_AQUUPY STANDARD; PRT; 231 AA.

ID USF_AQUUPY STANDARD; PRT; 231 AA.
 AC P46209;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 01-OCT-1996 (Rel. 34, Last annotation update)
 DE usf protein.
 GN USF.
 OS Aquifex pyrophilus.
 OC Bacteria; Aquificae; Aquificales; Aquificaceae; Aquifex.
 OX NCBI_TaxID=2714;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=KOLSA / DSM 6858;
 RX MEDLINE=96062250; PubMed=7592443;
 RA Behammer W., Shao Z., Mages W., Rachel R., Stetter K.O.,
 RA Schmitt R.;
 RT "Flagellar structure and hyperthermophily: analysis of a single
 RT flagellin gene and its product in Aquifex pyrophilus.";
 RJ J. Bacteriol. 177:6630-6637(1995).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; U17575; AAA88922.1; -
 DR InterPro: IPR002925; DLH.
 DR InterPro: IPR000379; Ser_estrs_site.
 DR Pfam; PF01738; DLH; 1.
 SQ SEQUENCE 231 AA; 26140 MW; C55E408839920195 CRC64;

Query Match 41.1%; Score 46; DB 1; Length 231;
 Best Local Similarity 53.3%; Pred. No. 3.7;
 Matches 8; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

OY 8 GOLLVFRANGVQYKG 22
 ||::| ||::|
 Db 2 GQVVEFEKNGVQVKG 16

Search completed: November 13, 2003, 09:45:29
 Job time : 9.92417 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 13, 2003, 09:31:40 ; Search time 33.1564 Seconds
(without alignments)
171.224 Million cell updates/sec

Title: US-09-290-049a-3
Perfect score: 112
Sequence: 1 TGARTINGQILYFRANGVQVKG 22

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues
Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

SPTREMBL_23:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	90.4	591	2	Q8VUH3	Q8VUH3 streptococc
2	87	77.7	1590	Q59983	Q59983 streptococc
3	82	73.2	1599	Q00599	Q00599 streptococc
4	81	72.3	2835	Q8G9Q2	Q8G9Q2 leucococc
5	80	71.4	1518	Q00600	Q00600 streptococc
6	76	67.9	1590	Q55263	Q55263 streptococc
7	75	67.0	726	Q8DUM9	Q8DUM9 streptococc
8	75	67.0	1290	Q48756	Q48756 leucococc
9	75	67.0	1338	Q9WUX4	Q9WUX4 streptococc
10	73	65.2	1449	Q55264	Q55264 streptococc
11	72	64.3	1477	Q91466	Q91466 leucococc
12	72	64.3	1508	Q9EZH5	Q9EZH5 leucococc
13	72	64.3	1508	Q52224	Q52224 leucococc
14	72	64.3	1577	Q54178	Q54178 streptococc
15	68	60.7	563	Q54447	Q54447 streptococc
16	68	60.7	565	Q8DRV2	Q8DRV2 streptococc

17	68	60.7	1449	2	Q68542	Q68542 streptococc
18	66	58.9	1577	2	Q55265	Q55265 streptococc
19	65.5	58.5	330	2	Q55228	Q55228 streptococc
20	65	58.0	2057	2	Q9RE05	Q9RE05 leucococc
21	64	57.1	1512	2	Q9WUX5	Q9WUX5 streptococc
22	60.5	54.0	1554	2	Q8KZL5	Q8KZL5 streptococc
23	57	50.9	1527	2	Q9ZARA	Q9ZARA leucococc
24	57	50.9	1527	2	Q8KRE1	Q8KRE1 leucococc
25	57	50.9	1575	2	Q9LCH3	Q9LCH3 streptococc
26	55	49.1	2817	16	Q97K42	Q97K42 clostridium
27	54	48.2	221	5	Q44353	Q44353 caenorhabdi
28	53	47.3	1002	10	Q8GZY5	Q8GZY5 oryza sativ
29	50	44.6	831	2	Q50076	Q50076 plasmodium
30	49	43.8	372	5	Q8TW25	Q8TW25 plasmodium
31	49	43.8	491	16	Q97JB2	Q97JB2 clostridium
32	47	42.0	126	2	Q47171	Q47171 escherichia
33	47	42.0	126	2	Q47173	Q47173 escherichia
34	47	42.0	126	2	Q47172	Q47172 escherichia
35	47	42.0	255	2	Q66378	Q66378 clostridium
36	47	42.0	396	2	Q9P932	Q9P932 clostridium
37	47	42.0	559	11	Q8K4R0	Q8K4R0 ratius norv
38	47	42.0	781	11	Q89090	Q89090 mus musculu
39	47	42.0	784	11	Q89087	Q89087 mus musculu
40	46	41.1	608	2	Q8VQ55	Q8VQ55 streptococc
41	46	41.1	619	2	Q54972	Q54972 streptococc
42	46	41.1	619	16	Q8DR10	Q8DR10 streptococc
43	46	41.1	744	16	Q97T39	Q97T39 streptococc
44	46	41.1	1301	2	Q06307	Q06307 alicyclobac
45	46	41.1	2364	2	Q46342	Q46342 clostridium

ALIGNMENTS

RESULT 1

Q8VUH3 PRELIMINARY; PRT; 591 AA.
AC Q8VUH3;
DT 01-MAR-2002 (TREMBLrel. 20, Created)
DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
DT 01-OCT-2002 (TREMBLrel. 22, Last annotation update)
DE Glucosyltransferase SI (Fragment).
GN GTFC.
OS Streptococcus mutans.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1309;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Xc; Yamashita Y., Shibata Y., Nakano Y., Koga T.,
RA Tada H., Involved in Bacitracin Resistance in Streptococcus mutans.",
RT Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
RL EMBL: AB078507; BAB83942.1; -
DR EMBL: AB078507; BAB83942.1; -
DR InterPro: IPR002479; CM binding.
DR InterPro: IPR003318; Glyco_hydro_70.
DR Pfam: PF01473; CM binding_1; 9.
DR Pfam: PF02324; Glyco_hydro_70; 1.
DR Transferase.
KM NON TER
FT NON TER
SQ SEQUENCE 591 AA; 67094 MW; 0933DCB4421DAF30 CRC64;
Query Match 80.4%; Score 90; DB 2; Length 591;
Best local Similarity 72.7%; Pred. No. 4.6e-06;
Matches 16; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 TGARTINGQILYFRANGVQVKG 22
DB 465 TGARTVNGQRLYFRANGVQAKG 486

RESULT 2
Q59983

ID 059983 PRELIMINARY; PRT; 1590 AA.
 AC 059983;
 DT 01-NOV-1996 (TRENBLER. 01, Created)
 DT 01-NOV-1996 (TRENBLER. 01, Last sequence update)
 DT 01-OCT-2002 (TRENBLER. 22, Last annotation update)
 DE Glucosyltransferase-I precursor (EC 2.4.1.5).
 GN GTFI.
 OS Streptococcus sobrinus.
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 CC Streptococcus.
 NX NCBI_TaxID=1310;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=OM2176;
 RA MEDLINE=94146405; PubMed=8112602;
 RA Sato S., Inoue M., Handa N., Aizawa Y., Isebe Y., Katayama T.;
 RT "DNA sequence of the glucosyltransferase gene of serotype d
 RT Streptococcus sobrinus.";
 RL DNA Seq. 4:19-27(1993).
 DR EMBL; D13858; BAA02976.1; -;
 DR InterPro; IPR002479; CM binding.
 DR InterPro; IPR003318; Glyco_hydro_70.
 DR Pfam; PF01473; CM_binding_1; 16.
 DR Pfam; PF02324; Glyco_hydro_70; 1.
 DR Glycosyltransferase; Signal; Transferase.
 FT SIGNAL 1 38 POTENTIAL.
 FT CHAIN 1 1590 GLUCOSYLTRANSFERASE-I.
 SQ SEQUENCE 1590 AA; 175955 MW; C3C83A57CF3C2B0E CRC64;

Query Match 77.7%; Score 87; DB 2; Length 1590;
 Best Local Similarity 77.3%; Pred. No. 4.4e-05;
 Matches 17; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 TGARTINGOLLYFRANGVQVKG 22
 |||:|||||:|||||:|||||
 Db 1296 TGAQTIRGQKLYFKANGQVQVKG 1317

RESULT 3
 ID 000599 PRELIMINARY; PRT; 1599 AA.
 AC 000599;
 DT 01-NOV-1996 (TRENBLER. 01, Created)
 DT 01-NOV-1996 (TRENBLER. 01, Last sequence update)
 DT 01-OCT-2002 (TRENBLER. 22, Last annotation update)
 DE Glucosyltransferase S precursor (EC 2.4.1.5) (GTF) (Dextranucrase)
 DE (Sucrose 6-glucosyltransferase).
 GN GTFK.
 OS Streptococcus salivarius.
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 CC Streptococcus.
 NX NCBI_TaxID=1304;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 25975;
 RA MEDLINE=93381463; PubMed=8371114;
 RA Giffard P.M., Allen D.M., Milward C.P., Simpson C.L., Jacques N.A.;
 RT "Sequence of the gtfK gene of Streptococcus salivarius ATCC 25975 and
 RT evolution of the gtf genes of oral streptococci.";
 RL J. Gen. Microbiol. 139:1511-1522(1993).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 25975;
 RA MEDLINE=92148377; PubMed=1838391;
 RA Giffard P.M., Simpson C.L., Milward C.P., Jacques N.A.;
 RT "Molecular characterization of a cluster of at least two
 RT glucosyltransferase genes in Streptococcus salivarius ATCC 25975.";
 RL J. Gen. Microbiol. 137:2577-2593(1991).
 CC -1- FUNCTION: PRODUCTION OF EXTRACELLULAR GLUCANS, THAT ARE THOUGHT TO
 CC PLAY A KEY ROLE IN THE DEVELOPMENT OF THE DENTAL PLAQUE BECAUSE OF
 CC AGGREGATION OF BACTERIAL CELLS AND FOOD DEBRIS.
 CC -1- CATALYTIC ACTIVITY: SUCROSE + ((1,6)-ALPHA-D-GLUCOSYL) (N) = D-

CC FRUCTOSE + ((1,6)-ALPHA-D-GLUCOSYL) (N+1).
 CC -1- SUBCELLULAR LOCATION: EXTRACELLULAR.
 CC -1- DISEASE: DENTAL CARIES.
 CC -1- SIMILARITY: TO REGIONS OF BARLEY AND BACILLUS AMYLOLIQUEPACIENS
 CC ALPHA AMYLASES AND RABBIT GLYCOGEN PHOSPHORYLASE.
 DR EMBL; Z11872; CAA77898.1; -;
 DR EMBL; Z11873; CAA77901.1; -;
 DR EMBL; M61113; AAA26897.1; -;
 DR InterPro; IPR002479; CM_binding.
 DR InterPro; IPR003318; Glyco_hydro_70.
 DR Pfam; PF01473; CM_binding_1; 14.
 DR Pfam; PF02324; Glyco_hydro_70; 1.
 DR Transferase; Glycosyltransferase; Signal; Repeat; Dental caries.
 FT SIGNAL 1 42 POTENTIAL.
 FT CHAIN 1 1599 GLUCOSYLTRANSFERASE S.
 SQ SEQUENCE 1599 AA; 176480 MW; 24B77869E152B707 CRC64;

QY 1 TGARTINGOLLYFRANGVQVKG 22
 |||:|||||:|||||:|||||
 Db 1402 TGAQVINGQHLVFDANGVQVKG 1423

RESULT 4
 ID 08G9Q2 PRELIMINARY; PRT; 2835 AA.
 AC 08G9Q2;
 DT 01-MAR-2003 (TRENBLER. 23, Created)
 DT 01-MAR-2003 (TRENBLER. 23, Last sequence update)
 DT 01-MAR-2003 (TRENBLER. 23, Last annotation update)
 DE Dextranucrase (EC 2.4.1.5) (Fragment).
 GN DSRB.
 OS Leuconostoc mesenteroides.
 OC Bacteria; Firmicutes; Lactobacillales; Leuconostoc.
 NX NCBI_TaxID=1245;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC MEDLINE=22231661; PubMed=12270834;
 RA Bozonnet S., Dols-Lafargue M., Fabre E., Pizzut S., Renaud-Simeon M.,
 RA Mondan P., Willemot R.M.;
 RT "Molecular characterization of DSR-E, an alpha-1,2 linkage
 RT synthesising dextranucrase with two catalytic domains.";
 RL J. Bacteriol. 184:5753-5761(2002).
 DR EMBL; AJ430204; CAD22883.1; -;
 DR Transferase; Glycosyltransferase.
 FT NON TER 1 1
 SQ SEQUENCE 2835 AA; 313264 MW; D03262CD735399D CRC64;

Query Match 72.3%; Score 81; DB 2; Length 2835;
 Best Local Similarity 72.7%; Pred. No. 0.00082;
 Matches 16; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 TGARTINGOLLYFRANGVQVKG 22
 |||:|||||:|||||:|||||
 Db 1644 TGAQTIRGQKLYFKANGQVQVKG 1665

RESULT 5
 ID 000600 PRELIMINARY; PRT; 1518 AA.
 AC 000600;
 DT 01-NOV-1996 (TRENBLER. 01, Created)
 DT 01-NOV-1996 (TRENBLER. 01, Last sequence update)
 DT 01-OCT-2002 (TRENBLER. 22, Last annotation update)
 DE Glucosyltransferase I (EC 2.4.1.5) (GTF) (Dextranucrase) (Sucrose 6-
 DE glucosyltransferase).
 GN GTFU.
 OS Streptococcus salivarius.
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 CC Streptococcus.

OX NCBI_TaxID=1304;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 25975;
 RX MEDLINE=92148377; PubMed=1838391;
 RA Gifford P.M., Simpson C.L., Milward C.P., Jacques N.A.;
 RT "Molecular characterization of a cluster of at least two
 RT glucosyltransferase genes in Streptococcus salivarius ATCC 25975.";
 RL J. Gen. Microbiol. 137:2577-2593(1991).
 CC -1- FUNCTION: PRODUCTION OF EXTRACELLULAR GLUCANS, THAT ARE THOUGHT TO
 CC PLAY A KEY ROLE IN THE DEVELOPMENT OF THE DENTAL PLAQUE BECAUSE OF
 CC THEIR ABILITY TO ADHERE TO SMOOTH SURFACES AND MEDIATE THE
 CC AGGREGATION OF BACTERIAL CELLS AND FOOD DEBRIS.
 CC -1- CATALYTIC ACTIVITY: SUCROSE + ((1,6)-ALPHA-D-GLUCOSYL) (N) = D-
 CC FRUCTOSE + ((1,6)-ALPHA-D-GLUCOSYL) (N+1).
 CC -1- SUBCELLULAR LOCATION: EXTRACELLULAR.
 CC -1- DISEASE: DENTAL CARIES.
 DR EMBL: Z11873; CAA77900.1; -;
 DR EMBL: M6411; AAA26896.1; -;
 DR InterPro: IPR002479; CW binding.
 DR InterPro: IPR003318; Glyco_hydro_70.
 DR Pfam: PF01473; CW binding 1; 13.
 DR Pfam: PF02324; Glyco_hydro_70; 1.
 KM Transferrase: Glycosyltransferase; Repeat; Dental caries.
 FT DOMAIN 1307 1482
 FT REPEAT 1307 1338
 FT REPEAT 1339 1352
 FT REPEAT 1372 1403
 FT REPEAT 1404 1417
 FT REPEAT 1437 1468
 FT REPEAT 1469 1482
 FT REPEAT 6.
 SQ SEQUENCE 1518 AA; 167730 MW; DAA41F717098B59A CRC64;
 Query March 71.4%; Score 80; DB 2; Length 1518;
 Best Local Similarity 72.7%; Pred. No. 0.0058;
 Matches 16; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
 QY 1 TGARTINGQLLYFRANGVQVK 22
 DB 1383 TGAQVINGQKLYFNADGSGVKG 1404
 RESULT 6
 Q55263 PRELIMINARY; PRT; 1590 AA.
 ID Q55263;
 AC Q55263;
 DT 01-NOV-1996 (TREMblrel. 01, Created)
 DT 01-JAN-1998 (TREMblrel. 05, Last sequence update)
 DT 01-OCT-2002 (TREMblrel. 22, Last annotation update)
 DE GTF-1.
 GN GLUCOSYLTRANSFERASE.
 OS Streptococcus sobrinus.
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 OX NCBI_TaxID=1310;
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 33478;
 RA Sato S.;
 RT "DNA and amino-acid sequences of water-insoluble-glycan synthetase
 RT produced from Streptococcus sobrinus ATCC 33478.";
 RL Ann. Kagoshima Univ. Dental School 16:23-29(1996).
 DR EMBL: D63570; BAA09792.1; -;
 DR InterPro: IPR002479; CW binding.
 DR InterPro: IPR003318; Glyco_hydro_70.
 DR Pfam: PF01473; CW binding 1; 15.
 DR Pfam: PF02324; Glyco_hydro_70; 1.
 SQ SEQUENCE 1590 AA; 176058 MW; 9DF7A3F2C6B4FD43 CRC64;
 Query March 67.9%; Score 76; DB 2; Length 1590;
 Best Local Similarity 68.2%; Pred. No. 0.0028;
 Matches 15; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 1 TGARTINGQLLYFRANGVQVK 22
 DB 1231 TGAQTVGKQKLYFNANGSGVKG 1252
 RESULT 7
 Q8DUW9 PRELIMINARY; PRT; 726 AA.
 ID Q8DUW9;
 AC Q8DUW9;
 DT 01-MAR-2003 (TREMblrel. 23, Created)
 DT 01-MAR-2003 (TREMblrel. 23, Last sequence update)
 DT 01-MAR-2003 (TREMblrel. 23, Last annotation update)
 DE Putative glycan-binding protein D, BglB-like protein.
 GN GBP OR SMU.772.
 OS Streptococcus mutans.
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 OX NCBI_TaxID=1309;
 RP SEQUENCE FROM N.A.
 RC STRAIN=UAI59 / ATCC 700610 / Serotype C;
 RX MEDLINE=22285063; PubMed=12397186;
 RA Ajdic D., Moshan W.M., McLaughlin R.E., Savic G., Chang J.,
 RA Carson M.B., Primeaux C., Tian R., Xenon S., Jia H., Lin S., Qian Y.,
 RA Li S., Zhu H., Najjar F., Lai H., White J., Roe B.A., Ferretti J.J.;
 RT "Genome sequence of Streptococcus mutans UAI59, a cariogenic dental
 RT pathogen.";
 RT Proc. Natl. Acad. Sci. U.S.A. 99:14434-14439(2002).
 RL EMBL: AB014918; BAA58492.1; -;
 KM Complete proteome.
 SQ SEQUENCE 726 AA; 79787 MW; B2AB7C06F6E7CB4E CRC64;
 Query March 67.0%; Score 75; DB 16; Length 726;
 Best Local Similarity 63.6%; Pred. No. 0.0016;
 Matches 14; Conservative 4; Mismatches 4; Indels 0; Gaps 0;
 QY 1 TGARTINGQLLYFRANGVQVK 22
 DB 235 TGSITVNGQSLYFNDSGSGVKG 256
 RESULT 8
 Q48756 PRELIMINARY; PRT; 1290 AA.
 ID Q48756;
 AC Q48756;
 DT 01-NOV-1996 (TREMblrel. 01, Created)
 DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)
 DT 01-OCT-2002 (TREMblrel. 22, Last annotation update)
 DE Dextranucrase.
 GN Leuconostoc mesenteroides.
 OS Bacteria; Firmicutes; Lactobacillales; Leuconostoc.
 OX NCBI_TaxID=1245;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=NRRL B1299;
 RX MEDLINE=97136686; PubMed=8982063;
 RA Monchois V., Willemot R.M., Renaud-Simeon M., Croux C., Monsan P.;
 RT "Cloning and sequencing of a gene coding for a novel dextranucrase
 RT from Leuconostoc mesenteroides NRRL B-1299 synthesizing only alpha (1-
 RT 6) and alpha (1-3) linkages.";
 RL Gene 182:23-32(1996).
 DR EMBL: U38181; AAA40875.1; -;
 DR InterPro: IPR002479; CW binding.
 DR InterPro: IPR003318; Glyco_hydro_70.
 DR Pfam: PF01473; CW binding 1; 11.
 DR Pfam: PF02324; Glyco_hydro_70; 1.
 SQ SEQUENCE 1290 AA; 145590 MW; 3555C2B96B749F9AA CRC64;
 Query March 67.0%; Score 75; DB 2; Length 1290;
 Best Local Similarity 68.2%; Pred. No. 0.0032;
 Matches 15; Conservative 2; Mismatches 5; Indels 0; Gaps 0;
 QY 1 TGARTINGQLLYFRANGVQVK 22

Db 1245 TGQVINGQTLTYPDAGRQVKG 1266

RESULT 9

Q9WXJ4 PRELIMINARY; PRT; 1338 AA.

AC 09WXJ4
 DT 01-NOV-1999 (TREMBlrel. 12, Created)
 DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
 DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
 DE GTF-S.
 GN GTF.
 OS Streptococcus criceti.
 OC Plasmid pAM1.
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 OX NCBI_TaxID=1333;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=HS-6;
 RA Inoue M., Fukui K., Miyagi A.;
 RT "S. cricetus glucosyltransferase (gifs and gftf) genes."
 RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AB026123; BAA77236.1; -
 DR InterPro: IPR002479; CW binding.
 DR InterPro: IPR003318; Glyco_hydro_70.
 DR Pfam: PF01473; CW binding_1; 10.
 DR Pfam: PF02324; Glyco_hydro_70; 1.
 DR Plasmid.
 KW
 SQ SEQUENCE 1338 AA; 148558 MW; 0A90CBE10E15D99B CRC64;

Query Match 67.0%; Score 75; DB 2; Length 1338;
 Best Local Similarity 63.6%; Pred. No. 0.0033;
 Matches 14; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

QY 1 TGARTINGQLTYFRANGVQVKG 22
 Db 1207 TGSQTAGQXVFQPNQVQVKG 1228

RESULT 10

Q55264 PRELIMINARY; PRT; 1449 AA.

AC 055264
 DT 01-NOV-1996 (TREMBlrel. 01, Created)
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
 DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
 DE Glucosyltransferase precursor.
 GN GTF.
 OS Streptococcus salivarius.
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 OX NCBI_TaxID=1304;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC MEDLINE=95122197; PubMed=7822030;
 RA Simpson C.L., Giffard P.M., Jacques N.A.;
 RT "Streptococcus salivarius ATCC 25975 possesses at least two genes
 coding for primer-independent glucosyltransferases."
 RL Infect. Immun. 63:609-621(1995).
 DR EMBL: L35495; AAC41412.1; -
 DR InterPro: IPR002479; CW binding.
 DR InterPro: IPR003318; Glyco_hydro_70.
 DR Pfam: PF01473; CW binding_1; 8.
 DR Pfam: PF02324; Glyco_hydro_70; 1.
 DR Signal; Transferase.
 KM SIGNAL
 FT SIGNAL 1 35
 FT CHAIN 36 1449
 FT SEQUENCE 1449 AA; 159984 MW; D62F07306E86A46 CRC64;

Query Match 65.2%; Score 73; DB 2; Length 1449;
 Best Local Similarity 68.2%; Pred. No. 0.0077;

Matches 15; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 1 TGARTINGQLTYFRANGVQVKG 22

Db 1258 TGHQNINGQELTFDNNQVQVKG 1279

RESULT 11

Q9L466 PRELIMINARY; PRT; 1477 AA.

AC 09L466
 DT 01-OCT-2000 (TREMBlrel. 15, Created)
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
 DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
 DE Dextranucrase (EC 2.4.1.5).
 GN DSR.
 OS Leuconostoc mesenteroides.
 OC Bacteria; Firmicutes; Lactobacillales; Leuconostoc.
 OX NCBI_TaxID=1245;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=NRRL B-1355;
 RA Arguello-Morales M.A., Renaud-Simeon K., Pizzut S., Sarcabal P.,
 RA Willemot R.M., Monsan P.;
 RT "Sequence analysis of the gene encoding alternansucrase, a sucrose
 glucosyltransferase from Leuconostoc mesenteroides NRRL B-1355."
 RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AJ250172; CAB76565.1; -
 DR InterPro: IPR002479; CW binding.
 DR InterPro: IPR003318; Glyco_hydro_70.
 DR Pfam: PF01473; CW binding_1; 14.
 DR Pfam: PF02324; Glyco_hydro_70; 1.
 DR Glycosyltransferase; Transferase.
 KW
 SQ SEQUENCE 1477 AA; 164887 MW; E6F5710DEDFC8B31 CRC64;

Query Match 64.3%; Score 72; DB 2; Length 1477;
 Best Local Similarity 68.2%; Pred. No. 0.011;
 Matches 15; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 1 TGARTINGQLTYFRANGVQVKG 22

Db 1368 TGLVNINGMLKXFQANGVQVKG 1389

RESULT 12

Q9EZHS PRELIMINARY; PRT; 1508 AA.

AC 09EZHS
 DT 01-MAR-2001 (TREMBlrel. 16, Created)
 DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
 DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
 DE Dextranucrase DsrB742.
 GN DSRB742.
 OS Leuconostoc mesenteroides.
 OC Bacteria; Firmicutes; Lactobacillales; Leuconostoc.
 OX NCBI_TaxID=1245;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=B-742CB;
 RA Kim H.-S., Kim D., Ryu H.-J., Robyt J.F.;
 RT "Leuconostoc mesenteroides B-742CB, a dextranucrase gene."
 RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AP294469; AAG38021.1; -
 DR InterPro: IPR002479; CW binding.
 DR InterPro: IPR003318; Glyco_hydro_70.
 DR Pfam: PF01473; CW binding_1; 14.
 DR Pfam: PF02324; Glyco_hydro_70; 1.
 DR SIGNAL
 SQ SEQUENCE 1508 AA; 168542 MW; E2FCEA0F97AE4F3A CRC64;

Query Match 64.3%; Score 72; DB 2; Length 1508;
 Best Local Similarity 68.2%; Pred. No. 0.012;
 Matches 15; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

```

Qy      1 TGARTINGQLLYFRANGVQYKG 22
Db      1399 TGLVNINGNLKTYFGANGROVKG 1420

RESULT 13
ID      052224      PRELIMINARY;      PRT; 1508 AA.
AC      052224;
DT      01-JUN-1998 (TREMBlrel. 06, Created)
DT      01-JUN-1998 (TREMBlrel. 06, Last sequence update)
DT      01-OCT-2002 (TREMBlrel. 22, Last annotation update)
DE      Glucosyltransferase (EC 2.4.1.5).
GN      DSRB.
OS      Leuconostoc mesenteroides.
OC      Bacteria; Firmicutes; Lactobacillales; Leuconostoc.
OX      NCBI_TaxID=1245;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=NRRL B-1299;
RA      Monchois V., Renaud-Simeon M., Monsan P., Willemot R.M.;
RT      "Cloning and sequencing of a gene coding for an extracellular
RT      dextranucrase (DSRB) from Leuconostoc mesenteroides NRRL B-1299
RT      synthesizing only a a(1-6) glucan."
RL      FEWS Microbiol. Lett. 0:0-0(1998).
DR      EMBL; AF030129; AAB95453.1; -
DR      InterPro; IPR002479; CW binding.
DR      InterPro; IPR003318; Glyco_hydro_70.
DR      Pfam; PF01473; CW_binding_1; 14.
DR      Pfam; PF02324; Glyco_hydro_70; 1.
KW      Glucosyltransferase; Transferase.
SQ      SEQUENCE 1508 AA; 16511 MW; E70CECB57A70D1F0 CRC64;

Query Match
Best Local Similarity 64.3%; Score 72; DB 2; Length 1508;
Matches 15; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

Qy      1 TGARTINGQLLYFRANGVQYKG 22
Db      1399 TGLVNINGNLKTYFGANGROVKG 1420

RESULT 14
ID      054178      PRELIMINARY;      PRT; 1577 AA.
AC      054178; OS4247;
DT      01-NOV-1996 (TREMBlrel. 01, Created)
DT      01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT      01-OCT-2002 (TREMBlrel. 22, Last annotation update)
DE      Glucosyltransferase.
GN      GTFG.
OS      Streptococcus gordonii Challis.
OC      Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC      Streptococcus.
OX      NCBI_TaxID=293390;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=CHALLIS;
RX      MEDLINE=96157084; PubMed=8586195;
RA      Vickerman M.M., Sulavik M.C., Clewell D.B.;
RT      "Molecular analysis of Streptococcus gordonii glucosyltransferase
RT      phase variants."
RL      Dev. Biol. Stand. 85:309-314(1995).
RN      [2]
RP      SEQUENCE OF 1-96 FROM N.A.
RC      STRAIN=CHALLIS;
RX      MEDLINE=92276337; PubMed=1534326;
RA      Sulavik M.C., Tardif G., Clewell D.B.;
RT      "Identification of a gene, rgg, which regulates expression of
RT      glucosyltransferase and influences the Spp phenotype of Streptococcus
RT      gordonii Challis."
RL      J. Bacteriol. 174:3577-3586(1992).
DR      EMBL; U12643; AAC3483.1; -

```

```

DR      EMBL; M89776; AAA26969.1; -
DR      InterPro; IPR002479; CW_binding.
DR      InterPro; IPR003318; Glyco_hydro_70.
DR      Pfam; PF01473; CW_binding_1; 18.
DR      Pfam; PF02324; Glyco_hydro_70; 1.
KW      Transferase.
SQ      SEQUENCE 1577 AA; 177805 MW; 5AE0328DC5E08D18 CRC64;

Query Match
Best Local Similarity 64.3%; Score 72; DB 2; Length 1577;
Matches 14; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

Qy      1 TGARTINGQLLYFRANGVQYKG 22
Db      1379 TGLKTTNNVLYPFDQKQYKG 1400

RESULT 15
ID      054447      PRELIMINARY;      PRT; 563 AA.
AC      054447;
DT      01-NOV-1996 (TREMBlrel. 01, Created)
DT      01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT      01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE      Glucan-binding protein (gbp).
DE      Streptococcus mutans.
OS      Streptococcus mutans.
OC      Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC      Streptococcus.
OX      NCBI_TaxID=1309;
RN      [1]
RP      SEQUENCE FROM N.A.
RA      Banas J.A., Russell R.R.B., Ferretti J.J.;
RT      "Sequence analysis of the gene for the glucan-binding protein of
RT      Streptococcus mutans INGBRITT."
RL      Submitted (MAY-1990) to the EMBL/genbank/DBJ databases.
DR      EMBL; M30945; AAA26894.1; -
DR      InterPro; IPR002479; CW_binding.
DR      Pfam; PF01473; CW_binding_1; 9.
SQ      SEQUENCE 563 AA; 62932 MW; A0BB6DAA42D1B5F CRC64;

Query Match
Best Local Similarity 60.7%; Score 68; DB 2; Length 563;
Matches 14; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

Qy      2 GARTINGQLLYFRANGVQYKG 22
Db      276 GWRITGGKYYFDITNGVQYKG 296

Search completed: November 13, 2003, 09:44:01
Job time : 35.1564 secs

```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Comugen Ltd.

OM protein - protein search, using sw model

Run on: November 13, 2003, 09:31:40 ; Search time 14.2844 Seconds
(without alignments)
65.165 Million cell updates/sec

Title: US-09-290-049a-3
Perfect score: 112
Sequence: 1 TGARTINGQLYFRANGVQVKG 22

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/ptodata/2/1aa/5A_COMB.pep:*
2: /cgn2_6/ptodata/2/1aa/5B_COMB.pep:*
3: /cgn2_6/ptodata/2/1aa/6A_COMB.pep:*
4: /cgn2_6/ptodata/2/1aa/6B_COMB.pep:*
5: /cgn2_6/ptodata/2/1aa/PTCTUS_COMB.pep:*
6: /cgn2_6/ptodata/2/1aa/backfill1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	112	100.0	1475	3	US-09-007-999-2
2	112	100.0	1475	3	US-09-210-361-2
3	112	100.0	1475	4	US-09-740-274-2
4	87	77.7	22	1	US-08-057-162B-3
5	78	69.6	1375	3	US-09-210-361-4
6	78	69.6	1375	4	US-09-740-274-4
7	67	59.8	1430	3	US-09-008-172-2
8	67	59.8	1430	3	US-09-210-361-6
9	67	59.8	1430	4	US-09-740-274-6
10	66	58.9	1577	2	US-08-793-824-2
11	65	58.0	2057	4	US-09-499-203-2
12	49.5	44.2	811	1	US-08-480-604A-7
13	49.5	44.2	811	2	US-08-405-496A-7
14	49.5	44.2	811	3	US-08-915-136-7
15	49.5	44.2	811	4	US-08-957-310-7
16	49.5	44.2	811	4	US-10-011-366-7
17	49.5	44.2	812	1	US-08-480-604A-29
18	49.5	44.2	812	3	US-08-915-136-23
19	49.5	44.2	812	3	US-08-480-604A-6
20	49.5	44.2	2710	2	US-08-405-496A-6
21	49.5	44.2	2710	3	US-08-915-136-6
22	49.5	44.2	2710	4	US-08-957-310-6
23	49.5	44.2	2710	4	US-10-011-366-6
24	46	41.1	619	1	US-08-465-746-2
25	46	41.1	619	2	US-08-214-164-2
26	46	41.1	619	2	US-08-467-852A-3
27	46	41.1	619	2	US-08-246-636-2

28	46	41.1	619	2	US-08-247-491A-3	Sequence 3, Appl
29	46	41.1	619	2	US-08-319-795-2	Sequence 2, Appl
30	46	41.1	619	2	US-08-468-985-2	Sequence 2, Appl
31	46	41.1	619	3	US-08-312-945-2	Sequence 2, Appl
32	46	41.1	641	3	US-08-961-083-160	Sequence 160, App
33	46	41.1	641	4	US-09-536-784-160	Sequence 2, Appl
34	46	41.1	648	1	US-08-072-070-2	Sequence 2, Appl
35	46	41.1	648	1	US-08-469-434-2	Sequence 2, Appl
36	46	41.1	648	1	US-08-214-222-2	Sequence 2, Appl
37	46	41.1	648	2	US-08-467-852A-2	Sequence 2, Appl
38	46	41.1	648	2	US-08-468-718-2	Sequence 2, Appl
39	46	41.1	648	3	US-08-247-491A-2	Sequence 3, Appl
40	46	41.1	648	3	US-08-446-201-3	Sequence 23, Appl
41	46	41.1	695	1	US-08-127-499A-23	Sequence 23, Appl
42	46	41.1	695	1	US-08-482-847-23	Sequence 41, Appl
43	46	41.1	1231	4	US-08-714-741-41	Sequence 2, Appl
44	43	38.4	153	4	US-09-287-070-2	Sequence 73, Appl
45	43	38.4	1036	4	US-09-206-942-73	

ALIGNMENTS

```

RESULT 1
US-09-007-999-2
; Sequence 2, Application US/09007999
; Patent No. 6087559
; GENERAL INFORMATION:
; APPLICANT: Nichols, Scott E.
; TITLE OF INVENTION: Substitutes for Modified Starch and
; FILE REFERENCE: 0356D
; CURRENT APPLICATION NUMBER: US/09/007,999
; EARLIER FILING DATE: 1998-01-16
; EARLIER APPLICATION NUMBER: 08/478,704
; EARLIER FILING DATE: 1995-06-07
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 1475
; TYPE: PRT
; ORGANISM: Streptococcus mutans
US-09-007-999-2

Query Match      100.0%; Score 112; DB 3; Length 1475;
Best Local Similarity 100.0%; Pred. No. 1.5e-10;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 TGARTINGQLYFRANGVQVKG 22
DB      1300 TGARTINGQLYFRANGVQVKG 1321

RESULT 2
US-09-210-361-2
; Sequence 2, Application US/09210361
; Patent No. 6284479
; GENERAL INFORMATION:
; APPLICANT: Nichols, Scott E.
; TITLE OF INVENTION: Substitutes for Modified Starches and
; FILE REFERENCE: 0357CR
; CURRENT APPLICATION NUMBER: US/09/210,361
; EARLIER FILING DATE: 1998-12-11
; EARLIER APPLICATION NUMBER: 09/007,999
; EARLIER FILING DATE: 1998-01-16
; EARLIER APPLICATION NUMBER: 08/478,704
; EARLIER FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 09/009,620
; EARLIER FILING DATE: 1998-01-20
; EARLIER APPLICATION NUMBER: 08/485,243
; EARLIER FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 09/008,172

```

```

; EARLIER FILING DATE: 1998-01-16
; EARLIER APPLICATION NUMBER: 08/482,711
; EARLIER FILING DATE: 1995-06-07
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 1475
; TYPE: PRT
; ORGANISM: Streptococcus mutans
; US-09-210-361-2

Query Match
Best Local Similarity 100.0%; Score 112; DB 3; Length 1475;
Pred. No. 1.5e-10;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGARTINGOLLYFRANGOVYKG 22
Db 1300 TGARTINGOLLYFRANGOVYKG 1321

RESULT 3
; Sequence 2, Application US/09740274
; Patent No. 6465203
; GENERAL INFORMATION:
; APPLICANT: Nichols, Scott E.
; TITLE OF INVENTION: Glucan-containing Compositions and Paper
; FILE REFERENCE: 0357CRD
; CURRENT APPLICATION NUMBER: US/09/740,274
; PRIOR FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: 09/210,361
; PRIOR FILING DATE: 1998-12-11
; PRIOR APPLICATION NUMBER: 09/007,999
; PRIOR FILING DATE: 1998-01-16
; PRIOR APPLICATION NUMBER: 08/478,704
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: 09/009,620
; PRIOR FILING DATE: 1998-01-20
; PRIOR APPLICATION NUMBER: 08/485,243
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: 09/008,172
; PRIOR FILING DATE: 1998-01-16
; PRIOR APPLICATION NUMBER: 08/482,711
; PRIOR FILING DATE: 1995-06-07
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 1475
; TYPE: PRT
; ORGANISM: Streptococcus mutans
; US-09-740-274-2

Query Match
Best Local Similarity 100.0%; Score 112; DB 4; Length 1475;
Pred. No. 1.5e-10;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGARTINGOLLYFRANGOVYKG 22
Db 1300 TGARTINGOLLYFRANGOVYKG 1321

RESULT 4
; Sequence 3, Application US/08057162B
; Patent No. 5686075
; GENERAL INFORMATION:
; APPLICANT: Taubman, Martin A.
; APPLICANT: Smith, Daniel J.
; TITLE OF INVENTION: SYNTHETIC PEPTIDE VACCINE FOR DENTAL CARIES
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
; STREET: Two Militia Drive
```

```

; CITY: Lexington
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02173
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/057,162B
; FILING DATE: 30-APR-1993
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/877,295
; FILING DATE: 01-MAY-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Wagner, Richard W.
; REGISTRATION NUMBER: 34,480
; REFERENCE/DOCKET NUMBER: FDC92-01A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 861-6240
; TELEFAX: (617) 861-9540
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 22 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-057-162B-3

Query Match
Best Local Similarity 77.7%; Score 87; DB 1; Length 22;
Pred. No. 1.9e-08;
Matches 17; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 TGARTINGOLLYFRANGOVYKG 22
Db 1 TGAQTINGOLLYFRANGOVYKG 22

RESULT 5
; Sequence 4, Application US/09210361
; Patent No. 6284479
; GENERAL INFORMATION:
; APPLICANT: Nichols, Scott E.
; TITLE OF INVENTION: Substitutes for Modified Starches and
; FILE REFERENCE: 0357CR
; CURRENT APPLICATION NUMBER: US/09/210,361
; CURRENT FILING DATE: 1998-12-11
; EARLIER APPLICATION NUMBER: 09/007,999
; EARLIER FILING DATE: 1998-01-16
; EARLIER APPLICATION NUMBER: 08/478,704
; EARLIER FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 09/009,620
; EARLIER FILING DATE: 1998-01-20
; EARLIER APPLICATION NUMBER: 08/485,243
; EARLIER FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 09/008,172
; EARLIER FILING DATE: 1998-01-16
; EARLIER APPLICATION NUMBER: 08/482,711
; EARLIER FILING DATE: 1995-06-07
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 1375
; TYPE: PRT
; ORGANISM: streptococcus mutans
; US-09-210-361-4

Query Match
69.6%; Score 78; DB 3; Length 1375;
```

Best Local Similarity 68.2%; Pred. No. 9.9e-05;
Matches 15; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 1 TGARTINGOLLYFRANGVQYKG 22

Db 1264 TGTVTNGORLYFKPENGVOAKG 1285

RESULT 6
US-09-740-274-4
; Sequence 4, Application US/09740274
; Patent No. 6465203
; GENERAL INFORMATION:
; APPLICANT: Nichols, Scott E.
; TITLE OF INVENTION: Glucan-containing Compositions and Paper
; FILE REFERENCE: 0357CRD
; CURRENT FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: 09/210,361
; PRIOR FILING DATE: 1998-12-11
; PRIOR APPLICATION NUMBER: 09/007,999
; PRIOR FILING DATE: 1998-01-16
; PRIOR APPLICATION NUMBER: 08/478,704
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: 09/009,620
; PRIOR FILING DATE: 1998-01-20
; PRIOR APPLICATION NUMBER: 08/485,243
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: 09/008,172
; PRIOR FILING DATE: 1998-01-16
; PRIOR APPLICATION NUMBER: 08/482,711
; PRIOR FILING DATE: 1995-06-07
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 1375
; TYPE: PRT
; ORGANISM: streptococcus mutans
US-09-740-274-4

Query Match 69.6%; Score 78; DB 4; Length 1375;
Best Local Similarity 68.2%; Pred. No. 9.9e-05;
Matches 15; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 1 TGARTINGOLLYFRANGVQYKG 22

Db 1264 TGTVTNGORLYFKPENGVOAKG 1285

RESULT 7
US-09-008-172-2
; Sequence 2, Application US/09008172
; Patent No. 6127602
; GENERAL INFORMATION:
; APPLICANT: Nichols, Scott E.
; TITLE OF INVENTION: Substitutes for Modified Starch and
; FILE REFERENCE: 035BD
; CURRENT FILING DATE: 1998-01-16
; PRIOR APPLICATION NUMBER: 08/482,711
; PRIOR FILING DATE: 1995-06-07
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 1430
; TYPE: PRT
; ORGANISM: Streptococcus mutans
US-09-008-172-2

Query Match 59.8%; Score 67; DB 3; Length 1430;
Best Local Similarity 59.1%; Pred. No. 0.0081;
Matches 13; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY 1 TGARTINGOLLYFRANGVQYKG 22

Db 1332 TGSQTIAKLYFASDQYKG 1353

RESULT 8
US-09-210-361-6
; Sequence 6, Application US/09210361
; Patent No. 6284479
; GENERAL INFORMATION:
; APPLICANT: Nichols, Scott E.
; TITLE OF INVENTION: Substitutes for Modified Starches and
; FILE REFERENCE: 0357CR
; CURRENT FILING DATE: 1998-12-11
; PRIOR APPLICATION NUMBER: 09/007,999
; PRIOR FILING DATE: 1998-01-16
; PRIOR APPLICATION NUMBER: 08/478,704
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: 09/009,620
; PRIOR FILING DATE: 1998-01-20
; PRIOR APPLICATION NUMBER: 08/485,243
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: 09/008,172
; PRIOR FILING DATE: 1998-01-16
; PRIOR APPLICATION NUMBER: 08/482,711
; PRIOR FILING DATE: 1995-06-07
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 6
; LENGTH: 1430
; TYPE: PRT
; ORGANISM: streptococcus mutans
US-09-210-361-6

Query Match 59.8%; Score 67; DB 3; Length 1430;
Best Local Similarity 59.1%; Pred. No. 0.0081;
Matches 13; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY 1 TGARTINGOLLYFRANGVQYKG 22

Db 1332 TGSQTIAKLYFASDQYKG 1353

RESULT 9
US-09-740-274-6
; Sequence 6, Application US/09740274
; Patent No. 6465203
; GENERAL INFORMATION:
; APPLICANT: Nichols, Scott E.
; TITLE OF INVENTION: Glucan-containing Compositions and Paper
; FILE REFERENCE: 0357CRD
; CURRENT FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: 09/210,361
; PRIOR FILING DATE: 1998-12-11
; PRIOR APPLICATION NUMBER: 09/007,999
; PRIOR FILING DATE: 1998-01-16
; PRIOR APPLICATION NUMBER: 08/478,704
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: 09/009,620
; PRIOR FILING DATE: 1998-01-20
; PRIOR APPLICATION NUMBER: 08/485,243
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: 09/008,172
; PRIOR FILING DATE: 1998-01-16
; PRIOR APPLICATION NUMBER: 08/482,711
; PRIOR FILING DATE: 1995-06-07
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 6

LENGTH: 1430
TYPE: PRT
ORGANISM: streptococcus mutans
US-09-740-274-6

Query Match 59.8%; Score 67; DB 4; Length 1430;
Best Local Similarity 59.1%; Pred. No. 0.0081;
Matches 13; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

OY 1 TGARTINGQLLYFRANGVQYKG 22
DB 1332 TGSQTIAKKTLYFASDQYKG 1353

RESULT 10
US-08-793-824-2
Sequence 2, Application US/08793824
Patent No. 5981838
GENERAL INFORMATION:
APPLICANT: SIMPSON, Christine Lynn
APPLICANT: Giffard, Philip Morrison
APPLICANT: Jacques, Nicholas Anthony
TITLE OF INVENTION: Genetic Manipulation of plants to
TITLE OF INVENTION: Increase stored Carbohydrates
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESSES:
ADDRESS: Griffith Hack & Co
STREET: Level 8, 168 Walker Street
CITY: No. 5981838th Sydney
STATE: New South Wales
COUNTRY: Australia
ZIP: 2060
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/793,824
FILING DATE:
CLASSIFICATION: 800
PRIOR APPLICATION NUMBER: AU PM7643
FILING DATE: 24-AUG-1994
TELECOMMUNICATION INFORMATION:
TELEPHONE: 61 2 9957 5944
TELEFAX: 61 2 957 6288
TELEX: 26547
INFORMATION FOR SEQ. ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1577 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: not relevant
MOLECULE TYPE: protein
ORIGINAL SOURCE:
ORGANISM: Streptococcus salivarius
US-08-793-824-2

Query Match 58.9%; Score 66; DB 2; Length 1577;
Best Local Similarity 54.5%; Pred. No. 0.014;
Matches 12; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

OY 1 TGARTINGQLLYFRANGVQYKG 22
DB 1469 TGLQINNKYVYFGSNGAQYKG 1490

RESULT 11
US-09-499-203-2
Sequence 2, Application US/09499203
Patent No. 6570065
GENERAL INFORMATION:

APPLICANT: KOSSMANN, Jens
APPLICANT: WELSH, Thomas
APPLICANT: OUNANZ, Martin
APPLICANT: KNUTH, Karola
TITLE OF INVENTION: Nucleic Acid Molecules Encoding Alternansucrase
FILE REFERENCE: 147-196P
CURRENT APPLICATION NUMBER: US/09/499,203
CURRENT FILING DATE: 2000-02-08
NUMBER OF SEQ ID NOS: 54
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 2
LENGTH: 2057
TYPE: PRT
ORGANISM: Leuconostoc mesenteroides
US-09-499-203-2

Query Match 58.0%; Score 65; DB 4; Length 2057;
Best Local Similarity 57.1%; Pred. No. 0.028;
Matches 12; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

OY 2 GARTINGQLLYFRANGVQYKG 22
DB 261 GLQTIIDNLTQYFNGQVQYKG 281

RESULT 12
US-08-480-604A-7
Sequence 7, Application US/08480604A
Patent No. 5736139
GENERAL INFORMATION:
APPLICANT: KINK, JOHN A.
APPLICANT: THALLEY, BRUCE S.
APPLICANT: PADHAY, NISHA V.
APPLICANT: FIRCH, JOSEPH R.
APPLICANT: STAFFORD, DOUGLAS C.
TITLE OF INVENTION: VACCINE AND ANTITOXIN FOR TREATMENT AND
TITLE OF INVENTION: PREVENTION OF C. DIFFICILE DISEASE
NUMBER OF SEQUENCES: 32
CORRESPONDENCE ADDRESSES:
ADDRESS: MEDLEN & CARROLL, LLP
STREET: 220 MONTGOMERY STREET, SUITE 2200
CITY: SAN FRANCISCO
STATE: CALIFORNIA
COUNTRY: UNITED STATES OF AMERICA
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/480,604A
FILING DATE: 07-JUN-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/422,711
FILING DATE: 14-APR-1995
PRIOR APPLICATION NUMBER:
APPLICATION NUMBER: US 08/405,496
FILING DATE: 16-MAR-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/329,154
FILING DATE: 25-OCT-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/161,907
FILING DATE: 02-DEC-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/985,321
FILING DATE: 04-DEC-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/429,791
FILING DATE: 31-OCT-1989
ATTORNEY/AGENT INFORMATION:

NAME: INGOLIA, DIANE E.
REGISTRATION NUMBER: 40,027
REFERENCE/DOCKET NUMBER: OPHD-01763
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 705-8410
TELEFAX: (415) 397-8338
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 811 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: protein
US-08-480-604A-7

Query Match 44.2%; Score 49.5; DB 1; Length 811;
Best Local Similarity 52.2%; Pred. No. 4.2;
Matches 12; Conservative 2; Mismatches 8; Indels 1; Gaps 1;

QY 1 TGARTINGQLLYFRAN-GVQVKG 22
Db 587 TGLRTIDGKKYENTNTAVAVTG 609

RESULT 13
US-08-405-496A-7
Sequence 7, Application US/08405496A
Patent No. 5919665
GENERAL INFORMATION:
APPLICANT: WILLIAMS, JAMES A.
TITLE OF INVENTION: VACCINE FOR CLOSTRIDIUM BOTULINUM
TITLE OF INVENTION: NEUROTOXIN
NUMBER OF SEQUENCES: 30
CORRESPONDENCE ADDRESS:
ADDRESSEE: MEDLEN & CARROLL, LLP
STREET: 220 MONTGOMERY STREET, SUITE 2200
CITY: SAN FRANCISCO
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94104

COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/405,496A
FILING DATE: 16-MAR-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/329,154
FILING DATE: 25-OCT-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/161,907
FILING DATE: 02-DEC-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/985,321
FILING DATE: 04-DEC-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/429,791
FILING DATE: 31-OCT-1989
ATTORNEY/AGENT INFORMATION:
NAME: INGOLIA, DIANE E.
REGISTRATION NUMBER: 40,027
REFERENCE/DOCKET NUMBER: OPHD-01308
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 705-8410
TELEFAX: (415) 397-8338
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 811 amino acids
TYPE: amino acid
STRANDEDNESS: unknown

TOPOLOGY: unknown
MOLECULE TYPE: protein
US-08-405-496A-7

Query Match 44.2%; Score 49.5; DB 2; Length 811;
Best Local Similarity 52.2%; Pred. No. 4.2;
Matches 12; Conservative 2; Mismatches 8; Indels 1; Gaps 1;

QY 1 TGARTINGQLLYFRAN-GVQVKG 22
Db 587 TGLRTIDGKKYENTNTAVAVTG 609

RESULT 14
US-08-915-136-7
Sequence 7, Application US/08915136
Patent No. 6290960
GENERAL INFORMATION:
APPLICANT: KINK, JOHN A.
APPLICANT: THALLEY, BRUCE S.
APPLICANT: PADHYE, NISHA V.
APPLICANT: FIRCA, JOSEPH R.
APPLICANT: STAFFORD, DOUGLAS C.
TITLE OF INVENTION: VACCINE AND ANTITOXIN FOR TREATMENT AND
TITLE OF INVENTION: PREVENTION OF C. DIFFICILE DISEASE
NUMBER OF SEQUENCES: 32
CORRESPONDENCE ADDRESS:
ADDRESSEE: MEDLEN & CARROLL, LLP
STREET: 220 MONTGOMERY STREET, SUITE 2200
CITY: SAN FRANCISCO
STATE: CALIFORNIA
COUNTRY: UNITED STATES OF AMERICA
ZIP: 94104

COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/915,136
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/480,604
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/405,496
FILING DATE: 16-MAR-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/329,154
FILING DATE: 25-OCT-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/161,907
FILING DATE: 02-DEC-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/985,321
FILING DATE: 04-DEC-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/429,791
FILING DATE: 31-OCT-1989
ATTORNEY/AGENT INFORMATION:
NAME: INGOLIA, DIANE E.
REGISTRATION NUMBER: 40,027
REFERENCE/DOCKET NUMBER: OPHD-01763
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 705-8410
TELEFAX: (415) 397-8338
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 811 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown

MOLECULE TYPE: protein
US-08-915-136-7

Query March 44.2%; Score 49.5; DB 3; Length 811;
Best Local Similarity 52.2%; Pred. No. 4.2;
Matches 12; Conservative 2; Mismatches 8; Indels 1; Gaps 1;

QY 1 TGARTINGOLLYFRAN-GVQYKG 22
Db 587 TGLRTIDGKKYFNTNTAVAVTG 609

RESULT 15
US-08-957-310-7

Sequence 7, Application US/08957310
Patent No. 6365158

GENERAL INFORMATION:

APPLICANT: Williams, James A.

TITLE OF INVENTION: IDENTIFICATION OF NEUTRALIZING EPITOPES

TITLE OF INVENTION: OF TOXIN A AND TOXIN B FOR THE TREATMENT OF C. DIFFICILE

TITLE OF INVENTION: DISEASE

NUMBER OF SEQUENCES: 22

CORRESPONDENCE ADDRESS:

ADDRESSES: Medlen & Carroll

STREET: 220 Montgomery Street, Suite 2200

CITY: San Francisco

STATE: California

COUNTRY: United States of America

ZIP: 94104

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/957,310

FILING DATE: 23-OCT-1997

CLASSIFICATION: 424

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/329,154

FILING DATE: 24-OCT-1994

APPLICATION NUMBER: US 08/161,907

FILING DATE: 02-DEC-1993

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/985,321

FILING DATE: 04-DEC-1992

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/429,791

FILING DATE: 31-OCT-1989

ATTORNEY/AGENT INFORMATION:

NAME: Ingolia, Diane E.

REGISTRATION NUMBER: 40,027

REFERENCE/DOCKET NUMBER: OPND-01121

TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 705-8410

TELEFAX: (415) 397-8338

INFORMATION FOR SEQ ID NO: 7:

SEQUENCE CHARACTERISTICS:

LENGTH: 811 amino acids

TYPE: amino acid

STRANDEDNESS: unknown

TOPOLOGY: unknown

MOLECULE TYPE: protein

US-08-957-310-7

Query Match 44.2%; Score 49.5; DB 4; Length 811;

Best Local Similarity 52.2%; Pred. No. 4.2;

Matches 12; Conservative 2; Mismatches 8; Indels 1; Gaps 1;

QY 1 TGARTINGOLLYFRAN-GVQYKG 22

Db 587 TGLRTIDGKKYFNTNTAVAVTG 609

Search completed: November 13, 2003, 09:47:58
Job time: 15.2844 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Comphen Ltd.

OM protein - protein search, using bw model

Run on: November 13, 2003, 09:45:40 / Search time 28.7773 Seconds
(without alignments)
139.565 Million cell updates/sec

Title: US-09-290-049a-3
Perfect score: 112
Sequence: 1 TCGARTINGQLTYFRANGVQVKG 22

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 666188 seqs, 182559486 residues

Total number of hits satisfying chosen parameters: 666188

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database:

Published Applications AA:
1: /cgn2_6/ptodata/2/pubppaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/2/pubppaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/2/pubppaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/2/pubppaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/2/pubppaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/2/pubppaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/2/pubppaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/2/pubppaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/2/pubppaa/US09_PUBCOMB.pep:*
10: /cgn2_6/ptodata/2/pubppaa/US09_PUBCOMB.pep:*
11: /cgn2_6/ptodata/2/pubppaa/US09_NEW_PUB.pep:*
12: /cgn2_6/ptodata/2/pubppaa/US10_PUBCOMB.pep:*
13: /cgn2_6/ptodata/2/pubppaa/US10_PUBCOMB.pep:*
14: /cgn2_6/ptodata/2/pubppaa/US10_PUBCOMB.pep:*
15: /cgn2_6/ptodata/2/pubppaa/US10_PUBCOMB.pep:*
16: /cgn2_6/ptodata/2/pubppaa/US60_NEW_PUB.pep:*
17: /cgn2_6/ptodata/2/pubppaa/US60_PUBCOMB.pep:*
18: /cgn2_6/ptodata/2/pubppaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	112	100.0	1475	9 US-09-740-274-2	Sequence 2, Appli
2	78	69.6	1375	9 US-09-740-274-4	Sequence 4, Appli
3	67	59.8	1430	9 US-09-740-274-6	Sequence 6, Appli
4	49.5	44.2	811	15 US-10-011-366-7	Sequence 7, Appli
5	49.5	44.2	866	16 US-10-222-038-2	Sequence 2, Appli
6	49.5	44.2	2710	15 US-10-011-366-6	Sequence 6, Appli
7	49	43.8	343	15 US-10-275-560-13	Sequence 13, Appli
8	48	42.9	120	15 US-10-156-761-13796	Sequence 13796, A
9	46	41.1	619	11 US-09-882-774-1	Sequence 1, Appli
10	46	41.1	641	9 US-09-765-272-160	Sequence 160, App
11	45	41.1	744	11 US-09-769-787-184	Sequence 184, App
12	44	39.3	115	9 US-09-764-869-970	Sequence 970, App
13	44	39.3	115	9 US-10-091-504-970	Sequence 970, App
14	43.5	38.8	196	10 US-09-738-626-4344	Sequence 4344, Ap
15	43	38.4	153	9 US-09-287-070-2	Sequence 2, Appli

16	43	38.4	153	15 US-10-243-977-2	Sequence 2, Appli
17	43	38.4	349	12 US-10-274-694-13	Sequence 13, Appli
18	43	38.4	1036	12 US-10-193-764-69	Sequence 69, Appli
19	43	38.4	1477	12 US-10-193-764-67	Sequence 67, Appli
20	43	38.4	1477	14 US-10-092-880-4	Sequence 4, Appli
21	42	37.5	165	15 US-10-156-761-14299	Sequence 14299, A
22	42	37.5	284	15 US-10-156-761-10979	Sequence 10979, A
23	42	37.5	929	11 US-09-298-523B-60	Sequence 60, Appli
24	41.5	37.1	472	12 US-10-029-386-33685	Sequence 33685, A
25	41.5	37.1	472	12 US-09-826-509-469	Sequence 469, App
26	41.5	37.1	472	15 US-10-225-567A-84	Sequence 49, Appli
27	41	36.6	412	12 US-10-301-997-49	Sequence 49, Appli
28	41	36.6	412	12 US-10-301-997-87	Sequence 87, Appli
29	40	36.6	678	10 US-09-738-626-4369	Sequence 4369, Ap
30	40	35.7	400	10 US-09-738-626-4358	Sequence 4358, Ap
31	40	35.7	499	15 US-10-156-761-11375	Sequence 11375, A
32	40	35.7	673	15 US-10-294-561-3	Sequence 3, Appli
33	39.5	35.3	251	12 US-10-366-686-6	Sequence 6, Appli
34	39	34.8	91	15 US-10-011-366-8	Sequence 8, Appli
35	39	34.8	179	9 US-09-287-070-1	Sequence 1, Appli
36	39	34.8	179	15 US-10-243-977-1	Sequence 1, Appli
37	39	34.8	232	15 US-10-198-070-34	Sequence 34, Appli
38	39	34.8	345	15 US-10-128-714-3396	Sequence 3396, Ap
39	39	34.8	356	15 US-10-128-714-8396	Sequence 8396, Ap
40	39	34.8	372	15 US-10-177-293-348	Sequence 348, App
41	39	34.8	453	9 US-09-765-272-38	Sequence 38, Appli
42	39	34.8	539	11 US-09-298-523B-54	Sequence 54, Appli
43	39	34.8	631	12 US-09-829-382-25	Sequence 25, Appli
44	39	34.8	655	11 US-09-056-019-2	Sequence 2, Appli
45	39	34.8	663	11 US-09-298-523B-58	Sequence 58, Appli

ALIGNMENTS

RESULT 1
US-09-740-274-2
Sequence 2, Application US/09740274
Patent No. US20020031826A1
GENERAL INFORMATION:
APPLICANT: Nichols, Scott E.
TITLE OF INVENTION: Glucan-containing Compositions and Paper
FILE REFERENCE: 0357CRD
CURRENT APPLICATION NUMBER: US/09/740, 274
PRIOR FILING DATE: 2000-12-19
CURRENT FILING DATE: 09/210,361
PRIOR APPLICATION NUMBER: 1998-12-11
PRIOR FILING DATE: 1998-12-11
PRIOR APPLICATION NUMBER: 09/007,999
PRIOR FILING DATE: 1998-01-16
PRIOR APPLICATION NUMBER: 08/478,704
PRIOR FILING DATE: 1995-06-07
PRIOR APPLICATION NUMBER: 09/009,620
PRIOR FILING DATE: 1998-01-20
PRIOR APPLICATION NUMBER: 08/485,243
PRIOR FILING DATE: 1995-06-07
PRIOR APPLICATION NUMBER: 09/008,172
PRIOR FILING DATE: 1998-01-16
PRIOR APPLICATION NUMBER: 08/482,711
PRIOR FILING DATE: 1995-06-07
NUMBER OF SEQ ID NOS: 6
SOFTWARE: FASTSEQ for Windows Version 3.0
SEQ ID NO 2
LENGTH: 1475
TYPE: PRT
ORGANISM: Streptococcus mutans
US-09-740-274-2

Query Match 100.0%; Score 112; DB 9; Length 1475;
Best Local Similarity 100.0%; Pred. No. 3.5e-10;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1 TCGARTINGQLTYFRANGVQVKG 22
|||||

Db 1300 TGAATNGQLLYFRANGVQYKG 1321

RESULT 2

US-09-740-274-4
 ; Sequence 4, Application US/09740274
 ; Patent No. US20020031826A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Nichols, Scott E.
 ; TITLE OF INVENTION: Glucan-containing Compositions and Paper
 ; FILE REFERENCE: 0357CRD
 ; CURRENT APPLICATION NUMBER: US/09/740,274
 ; CURRENT FILING DATE: 2000-12-19
 ; PRIOR APPLICATION NUMBER: 09/210,361
 ; PRIOR FILING DATE: 1998-12-11
 ; PRIOR APPLICATION NUMBER: 09/007,999
 ; PRIOR FILING DATE: 1998-01-16
 ; PRIOR APPLICATION NUMBER: 08/478,704
 ; PRIOR FILING DATE: 1995-06-07
 ; PRIOR APPLICATION NUMBER: 09/009,620
 ; PRIOR FILING DATE: 1998-01-20
 ; PRIOR APPLICATION NUMBER: 08/485,243
 ; PRIOR FILING DATE: 1995-06-07
 ; PRIOR APPLICATION NUMBER: 09/008,172
 ; PRIOR FILING DATE: 1998-01-16
 ; PRIOR APPLICATION NUMBER: 08/482,711
 ; PRIOR FILING DATE: 1995-06-07
 ; NUMBER OF SEQ ID NOS: 6
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 4
 ; LENGTH: 1375
 ; TYPE: PRT
 ; ORGANISM: streptococcus mutans
 ; US-09-740-274-4

Query Match 69.6%; Score 79; DB 9; Length 1375;
 Best Local Similarity 68.2%; Pred. No. 0.00022;
 Matches 15; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 1 TGAATNGQLLYFRANGVQYKG 22
 Db 1264 TGTVTNGQLLYFRANGVQYKG 1285

RESULT 3

US-09-740-274-6
 ; Sequence 6, Application US/09740274
 ; Patent No. US20020031826A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Nichols, Scott E.
 ; TITLE OF INVENTION: Glucan-containing Compositions and Paper
 ; FILE REFERENCE: 0357CRD
 ; CURRENT APPLICATION NUMBER: US/09/740,274
 ; CURRENT FILING DATE: 2000-12-19
 ; PRIOR APPLICATION NUMBER: 09/210,361
 ; PRIOR FILING DATE: 1998-12-11
 ; PRIOR APPLICATION NUMBER: 09/007,999
 ; PRIOR FILING DATE: 1998-01-16
 ; PRIOR APPLICATION NUMBER: 08/478,704
 ; PRIOR FILING DATE: 1995-06-07
 ; PRIOR APPLICATION NUMBER: 09/009,620
 ; PRIOR FILING DATE: 1998-01-20
 ; PRIOR APPLICATION NUMBER: 08/485,243
 ; PRIOR FILING DATE: 1995-06-07
 ; PRIOR APPLICATION NUMBER: 09/008,172
 ; PRIOR FILING DATE: 1998-01-16
 ; PRIOR APPLICATION NUMBER: 08/482,711
 ; PRIOR FILING DATE: 1995-06-07
 ; NUMBER OF SEQ ID NOS: 6
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 6
 ; LENGTH: 1430
 ; TYPE: PRT

; ORGANISM: streptococcus mutans
 ; US-09-740-274-6

Query Match 59.8%; Score 67; DB 9; Length 1430;
 Best Local Similarity 59.1%; Pred. No. 0.018;
 Matches 13; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY 1 TGAATNGQLLYFRANGVQYKG 22
 Db 1332 TGSQTLAGKLYFRASDQYKG 1353

RESULT 4

US-10-011-366-7
 ; Sequence 7, Application US/10011366
 ; Publication No. US20030054493A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Williams, James A.
 ; Klink, John A.
 ; TITLE OF INVENTION: IDENTIFICATION OF NEUTRALIZING EPITOPES
 ; OF TOXIN A AND TOXIN B FOR THE TREATMENT OF C. DIFFICILE
 ; DISEASE
 ; NUMBER OF SEQUENCES: 22
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESS: Medlen & Carroll
 ; STREET: 220 Montgomery Street, Suite 2200
 ; CITY: San Francisco
 ; STATE: California
 ; COUNTRY: United States of America
 ; ZIP: 94104
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patentin Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/10/011,366
 ; FILING DATE: 16-NO. US20030054493A1-2001
 ; CLASSIFICATION: <Unknown>
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/957,310
 ; FILING DATE: 23-OCT-1997
 ; APPLICATION NUMBER: US 08/329,154
 ; FILING DATE: 24-OCT-1994
 ; APPLICATION NUMBER: US 08/161,907
 ; FILING DATE: 02-DEC-1993
 ; APPLICATION NUMBER: US 07/985,321
 ; FILING DATE: 04-DEC-1992
 ; APPLICATION NUMBER: US 07/429,791
 ; FILING DATE: 31-OCT-1989
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Ingolia, Diane E.
 ; REGISTRATION NUMBER: 40,027
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (415) 705-8410
 ; TELEFAX: (415) 397-8338
 ; INFORMATION FOR SEQ ID NO: 7:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 811 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: unknown
 ; TOPOLOGY: unknown
 ; MOLECULE TYPE: protein
 ; SEQUENCE DESCRIPTION: SEQ ID NO: 7:

US-10-011-366-7

Query Match 44.2%; Score 49.5; DB 15; Length 811;
 Best Local Similarity 52.2%; Pred. No. 9;
 Matches 12; Conservative 2; Mismatches 8; Indels 1; Gaps 1;

QY 1 TGAATNGQLLYFRANGVQYKG 22


```

; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATORI, MASAHISA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 13796
; LENGTH: 120
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-13796

```

```

Query Match          42.9%; Score 48; DB 15; Length 120;
Best Local Similarity 40.9%; Pred. No. 1.7;
Matches 9; Conservative 7; Mismatches 6; Indels 0; Gaps 0;

```

```

QY 1 TGARTINGQLYFRANGVQVKG 22
    |||:|||||:|||||:
Db 32 SGAPVYTGKVLMSADGKQVWG 53

```

```

RESULT 9
US-09-882-774-1
; Sequence 1, Application US/09882774
; Publication No. US20030021795A1
; GENERAL INFORMATION:
; APPLICANT: Houston, Michael E.
; APPLICANT: Hodges, Robert
; TITLE OF INVENTION: Use of Coiled-Coil Structural Scaffold to Generate
; FILE REFERENCE: 003592-007
; CURRENT APPLICATION NUMBER: US/09/882,774
; CURRENT FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: US 60/211,892
; PRIOR FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: US 60/213,387
; PRIOR FILING DATE: 2000-06-23
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 619
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-882-774-1

```

```

Query Match          41.1%; Score 46; DB 11; Length 619;
Best Local Similarity 40.9%; Pred. No. 26;
Matches 9; Conservative 1; Mismatches 12; Indels 0; Gaps 0;

```

```

QY 1 TGARTINGQLYFRANGVQVKG 22
    |||:|||||:|||||:
Db 462 TGMAYNGSWYIYNANGAMATG 483

```

```

RESULT 10
US-09-765-272-160
; Sequence 160, Application US/09765272
; Patent No. US20020061545A1
; GENERAL INFORMATION:
; APPLICANT: Choi et. al.
; TITLE OF INVENTION: Streptococcus pneumoniae Antigens and Vaccines
; NUMBER OF SEQUENCES: 452
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland

```

```

; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4MB storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/765,272
; FILING DATE: 22-Jan-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/961,083
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Brookes, A. Anders
; REGISTRATION NUMBER: 36,373
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 160:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 641 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULAR TYPE: Protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 160:
US-09-765-272-160

```

```

Query Match          41.1%; Score 46; DB 9; Length 641;
Best Local Similarity 40.9%; Pred. No. 27;
Matches 9; Conservative 1; Mismatches 12; Indels 0; Gaps 0;

```

```

QY 1 TGARTINGQLYFRANGVQVKG 22
    |||:|||||:|||||:
Db 544 TGMAYNGSWYIYNANGAMATG 565

```

```

RESULT 11
US-09-769-787-184
; Sequence 184, Application US/09769787
; Publication No. US20030091577A1
; GENERAL INFORMATION:
; APPLICANT: Microbial Technics Limited
; APPLICANT: Gilbert, Christophe FG
; APPLICANT: Hansbro, Philip M
; TITLE OF INVENTION: Proteins
; FILE REFERENCE: PWC/P21129MO
; CURRENT APPLICATION NUMBER: US/09/769,787
; CURRENT FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: GB 9816337.1
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: US 60/125164
; PRIOR FILING DATE: 1998-03-19
; NUMBER OF SEQ ID NOS: 388
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 184
; LENGTH: 744
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-769-787-184

```

```

Query Match          41.1%; Score 46; DB 11; Length 744;
Best Local Similarity 40.9%; Pred. No. 33;
Matches 9; Conservative 1; Mismatches 12; Indels 0; Gaps 0;

```

```

QY 1 TGARTINGQLYFRANGVQVKG 22
    |||:|||||:|||||:
Db 647 TGMAYNGSWYIYNANGAMATG 668

```

```
RESULT 12
US-09-764-869-970
; Sequence 970, Application US/09764869
; Patent No. US20020061521A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC007
; CURRENT APPLICATION NUMBER: US/09/764,869
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 2442
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 970
; LENGTH: 115
; TYPE: PRT
; ORGANISM: Homo sapiens
; NAME/KEY: SITE
; LOCATION: (1)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (5)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (6)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (7)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (8)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (26)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (28)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (30)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (46)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (50)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (58)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (59)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-764-869-970

Query Match          39.3%; Score 44; DB 9; Length 115;
Best Local Similarity 70.0%; Pred. No. 7.6;
Matches 7; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 5 TINGOLLYER 14
   |::|::|::|
Db 63 TUNGELLEFR 72

RESULT 13
US-10-091-504-970
; Sequence 970, Application US/10091504
; Publication No. US2003005908A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC007C1
```

```
; CURRENT APPLICATION NUMBER: US/10/091,504
; CURRENT FILING DATE: 2002-03-07
; NUMBER OF SEQ ID NOS: 2442
; Prior Application removed - See File Wrapper or Palm
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 970
; LENGTH: 115
; TYPE: PRT
; ORGANISM: Homo sapiens
; NAME/KEY: misc_feature
; LOCATION: (1)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: misc_feature
; LOCATION: (5)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: misc_feature
; LOCATION: (6)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: misc_feature
; LOCATION: (7)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: misc_feature
; LOCATION: (8)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: misc_feature
; LOCATION: (26)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: misc_feature
; LOCATION: (28)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: misc_feature
; LOCATION: (30)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: misc_feature
; LOCATION: (46)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: misc_feature
; LOCATION: (50)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: misc_feature
; LOCATION: (58)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: misc_feature
; LOCATION: (59)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-10-091-504-970

Query Match          39.3%; Score 44; DB 15; Length 115;
Best Local Similarity 70.0%; Pred. No. 7.6;
Matches 7; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 5 TINGOLLYER 14
   |::|::|::|
Db 63 TUNGELLEFR 72

RESULT 14
US-09-738-626-4344
; Sequence 4344, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAOKO
; APPLICANT: SENOH, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
```


GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 13, 2003, 09:31:40 ; Search time 38.5166 Seconds
(without alignments)
86.541 Million cell updates/sec

Title: US-09-290-049a-10

Sequence: 1 ANDHLSILEAWSDNDPTPLD 21

Scoring table: BLOSUM62
Gapop 10.0 , Gapept 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

```

1: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
19: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
20: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
21: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
22: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*
23: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.*
24: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	119	100.0	1017	23	AAU79285
2	119	100.0	1475	23	AAU98027
3	119	100.0	1475	23	AAU98030
4	119	100.0	1475	23	AAU98031
5	119	100.0	1475	23	AAU98032
6	119	100.0	1475	23	AAU98033
7	119	100.0	1475	23	AAU98034
8	119	100.0	1475	23	AAU98035
9	119	100.0	1475	23	AAU98036

10	119	100.0	1475	23	AAU98037
11	119	100.0	1475	23	AAU98038
12	119	100.0	1475	23	AAU98039
13	119	100.0	1475	23	AAU98040
14	119	100.0	1475	23	AAU98041
15	110	92.4	1375	23	AAU98028
16	110	92.4	1375	23	AAU98028
17	108	90.8	1592	14	AA32925
18	75	63.0	1430	23	AAU98029
19	75	63.0	1430	23	AAU98041
20	75	63.0	1430	23	AAU98042
21	75	63.0	1430	23	AAU98043
22	75	63.0	1430	23	AAU98044
23	75	63.0	1430	23	AAU98045
24	71	59.7	1527	23	AAU80055
25	69	58.0	2835	23	ABB98574
26	66	55.5	12	23	ABB98641
27	63	52.9	1577	17	AAU91047
28	60	50.4	12	23	ABB98642
29	60	50.4	2057	21	AAU10657
30	59	49.6	12	23	ABB98643
31	53	44.5	12	23	ABB98644
32	53	44.5	12	23	ABB98646
33	51	42.9	486	22	AA336552
34	50	42.0	305	21	AAU98845
35	50	42.0	339	21	AAU98844
36	50	42.0	649	22	AAU62860
37	49	41.2	12	23	ABB98578
38	49	41.2	12	23	ABB98648
39	49	41.2	302	22	AAU38411
40	49	41.2	1781	23	AAU74519
41	48.5	40.8	401	22	AAU43544
42	48	40.3	855	23	ABB98573
43	47.5	39.9	534	21	AAU67414
44	47	39.5	195	21	AAU69941
45	47	39.5	338	15	AAU65965

ALIGNMENTS

RESULT 1	AAU79285	standard; Protein; 1017 AA.
XX	AAU79285	
AC	AAU79285	
XX	AAU79285	
DT	13-AUG-2002	(first entry)
DE	Streptococcus mutans monoclonal antibody-related protein #2.	
KW	Antibody; dental caries; water insoluble glucan synthetase;	
KM	anti-carries; glucosyl transferase-B; immunotherapy.	
OS	Streptococcus mutans.	
XX	JP2002114709-A.	
XX	16-APR-2002.	
XX	04-OCT-2000; 2000JP-0304889.	
XX	04-OCT-2000; 2000JP-0304889.	
XX	04-OCT-2000; 2000JP-0304889.	
XX	(UYN1-) UNIV NIPPON.	
XX	WPI; 2002-448101/48.	
DR	Anti-carries agent composed of a monoclonal antibody against an	
XX	inhibitory enzyme against water insoluble glucan synthetase of glucosyl	
PT	transferase-B (GTP-B) of Streptococcus mutans	
XX	Claim 4; Page 17-19; 28pp; Japanese.	

S. mutans glucosyl
S. mutans GTP mut
S. mutans glucosyl
S. mutans glucosyl
S. mutans glucosyl
Streptococcus mut
S. mutans glucosyl
Streptococcus mut
Glucosyltransferase
S. mutans glucosyl
S. mutans glucosyl
S. mutans glucosyl
S. mutans glucosyl
S. mutans glucosyl
Leuconostoc mesent
Dextran saccharase
Dextrane-saccharas
Alpha-D-glucosyltr
Dextrane-saccharas
L. mesenteroides a
Dextrane-saccharas
Dextrane-saccharas
Dextrane-saccharas
Salmonella typhi c
Lactobacillus reut
Human polypeptide
Dextran saccharase
Arabidopsis aldehy
Arabidopsis thalia
T. niyeum GAPDH.

XX The invention relates to a monoclonal antibody against dental caries and
 CC an anti-carries agent composed of a monoclonal antibody produced by
 CC Streptococcus mutans, particularly mouse-hybridoma MHP126 (FERM P-17566)
 CC or mouse-hybridoma MHP136 (FERM P-17567), against an enzyme having
 CC inhibitive activity against water insoluble glucan synthetase of glucosyl
 CC transferase-B. The monoclonal antibody specifically inhibits water
 CC insoluble glucan synthetase of Streptococcus mutans produced glucosyl
 CC transferase-B and is used in the immunotherapy of dental caries. This
 CC sequence represents a Streptococcus mutans monoclonal antibody-related
 CC protein.
 CC
 SQ Sequence 1017 AA;
 Query Match 100.0%; Score 119; DB 23; Length 1017;
 Best Local Similarity 100.0%; Pred. No. 3.5e-09;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 1 ANDHLSILEAMSDNDTPYLHD 21
 447 ANDHLSILEAMSDNDTPYLHD 467
 RESULT 2
 AAU98027
 ID AAU98027 standard; Protein; 1475 AA.
 AC AAU98027;
 DT 27-AUG-2002 (first entry)
 DE S. mutans glucosyltransferase GTFB.
 XX Glucosyltransferase; GTFB; transgenic plant; paper sizing;
 KW coating composition; glucan; starch; latex; thermoplastic molecule;
 KM amyloplast; vacuole; paper manufacture.
 XX
 OS Streptococcus mutans.
 XX
 PN US2002031826-A1.
 PD 14-MAR-2002.
 PF 19-DEC-2000; 2000US-0740274.
 XX
 PR 11-DEC-1998; 98US-0210361.
 PR 07-JUN-1995; 95US-0478704.
 PR 07-JUN-1995; 95US-0482711.
 PR 07-JUN-1995; 95US-0485243.
 PR 16-JAN-1998; 98US-0007999.
 PR 16-JAN-1998; 98US-0008172.
 PR 20-JAN-1998; 98US-0009620.
 XX
 PA (NICH//) NICHOLS S E.
 XX
 PI Nichols SE;
 XX
 DR WPI, 2002-414332/44.
 DR N-PSDB; ABR52938.
 XX
 PT Glucosyltransferase B or D protein useful for producing a glucan useful
 as substitutes for and additions to modified starch and latexes in
 paper manufacture, comprises mutations in specific positions -
 XX
 PS Disclosure: Page 21-25; 44pp; English.
 XX
 CC The invention of an isolated protein comprising a glucosyltransferase
 CC (GTF) B polypeptide having changes at position from 1448V, D457N,
 CC D567T, K1014T, D457N/D567T, D457N/D571K, D567T/D571K,
 CC D567T/D571K/K1014T, I448V/D457N/D571K/K779Q/K1014T,
 CC Y169A/Y170A/Y171A, and K779Q or a GTF D polypeptide having
 CC changes at positions from T589D, T589B, N471D, N471D/T589D, and
 CC N471D/T589B. Also included are a glucan produced by the GTF mutant,

CC an isolated polynucleotide which encodes P1 or P2, or its complementary
 CC polynucleotide, a ribonucleic acid sequence encoding the GTF mutant,
 CC an expression cassette comprising the polynucleotide operably linked to a
 CC promoter, a vector comprising the expression cassette, host cell
 CC introduced with the vector, a transgenic plant comprising the
 CC vector, a seed or tuber from the transgenic plant, a paper sizing and/or
 CC coating composition comprising the mutant GTF, wild type or, starch, a latex,
 CC with a gene encoding the mutant GTF, wild type or, starch, a latex,
 CC thermoplastic molecule or their combinations or glucan and starch where
 CC the glucan is produced in the amyloplast and/or vacuole or a maize line
 CC deficient in starch biosynthesis, transformed with a gene encoding a
 CC glucosyltransferase B or D enzyme, wild-type or mutant and a paper
 CC comprising the glucan (paper sizing/coating agent). The vector is useful
 CC for producing a glucan in a plant. The method comprises transforming a
 CC plant cell with the vector, growing the plant cell under plant growing
 CC conditions to produce a regenerated plant and inducing expression of the
 CC polynucleotide for a time sufficient to produce the glucan in the
 CC regenerated plant, where the vector contains a transit sequence from
 CC ribulose biphosphate carboxylase small subunit, waxy, brittle-1 and
 CC chlorophyll AB binding protein to produce a transgenic plant, and glucan
 CC is produced in the amyloplast of potato or the vacuole of sugar beet.
 CC Glucans are useful as substitutes for and additions to modified starch
 CC and latexes in paper manufacture. Unlike prior art techniques, which
 CC require input materials that produce chemical effluents, paper
 CC manufacture utilizing the glucan produced by GTF, which utilizes
 CC biologically produced input materials, is more cost-effective and
 CC environmentally friendly. Moreover, glucans also exhibit thermoplastic
 CC properties and impart gloss to the paper during coating step.
 CC The present sequence represents GTFB.
 CC
 SQ Sequence 1475 AA;
 Query Match 100.0%; Score 119; DB 23; Length 1475;
 Best Local Similarity 100.0%; Pred. No. 5.4e-09;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 1 ANDHLSILEAMSDNDTPYLHD 21
 481 ANDHLSILEAMSDNDTPYLHD 501
 Db
 RESULT 3
 AAU98030
 ID AAU98030 standard; Protein; 1475 AA.
 AC AAU98030;
 DT 27-AUG-2002 (first entry)
 DE S. mutans glucosyltransferase GTFB mutant I448V.
 XX
 KW Glucosyltransferase; GTFB; transgenic plant; paper sizing;
 KM coating composition; glucan; starch; latex; thermoplastic molecule;
 XX amyloplast; vacuole; paper manufacture; mutant; mutain.
 XX
 OS Streptococcus mutans.
 XX
 PI Synthetic.
 XX
 FT Key Location/Qualifiers
 FT Misc-difference 448 /note= "Wild-type Ile substituted by Val"
 FT
 XX
 PN US2002031826-A1.
 PD 14-MAR-2002.
 PF 19-DEC-2000; 2000US-0740274.
 XX
 PR 11-DEC-1998; 98US-0210361.
 PR 07-JUN-1995; 95US-0478704.
 PR 07-JUN-1995; 95US-0482711.
 PR 07-JUN-1995; 95US-0485243.
 PR 16-JAN-1998; 98US-0007999.
 PR 16-JAN-1998; 98US-0007999.

PR 16-JAN-1998; 98US-0008172.
 PR 20-JAN-1998; 98US-0009620.
 XX
 XX
 PA (NICH/) NICHOLS S E.
 PI Nichols SE;
 XX
 XX WPI; 2002-414332/44.
 DR
 XX
 PT Glucosyltransferase B or D protein useful for producing a glucan useful
 PT as substitutes for and additions to modified starch and latexes in
 PT paper manufacture, comprises mutations in specific positions
 XX
 XX
 PS Claim 36; Page -: 44pp; English.
 XX
 CC The invention an isolated protein comprising a glucosyltransferase
 CC (GTF) B polypeptide having changes at position from 1448V, D457N,
 CC D567T, K1014T, D457N/D567T, D457N/D571K, D567T/D571K,
 CC D567T/D571K/K1014T, T448V/D457N/D567T/D571K/K779Q/K1014T,
 CC Y169A/Y170A/Y171A, and K779Q or a GTF D polypeptide having
 CC changes at positions from T589D, T589E, N471D, N471D/T589D, and
 CC N471D/T589E. Also included are a glucan produced by the GTF mutant,
 CC an isolated polynucleotide which encodes P1 or P2, or its complementary
 CC polynucleotide, a ribonucleic acid sequence encoding the GTF mutant,
 CC an expression cassette comprising the polynucleotide operably linked to a
 CC promoter, a vector comprising the expression cassette, host cell
 CC introduced with the vector, a transgenic plant comprising the
 CC vector, a seed or tuber from the transgenic plant, a paper sizing and/or
 CC coating composition comprising a glucan produced in a plant transformed
 CC with a gene encoding the mutant GTF, wild type or, starch, a latex,
 CC thermoplastic molecule or their combinations or glucan and starch where
 CC the glucan is produced in the amyloplast and/or vacuole or a maize line
 CC deficient in starch biosynthesis, transformed with a gene encoding a
 CC glucosyltransferase B or D enzyme, wild-type or mutant and a paper
 CC comprising the glucan (paper sizing/coating agent). The vector is useful
 CC for producing a glucan in a plant. The method comprises transforming a
 CC plant cell with the vector, growing the plant cell under plant growing
 CC conditions to produce a regenerated plant and inducing expression of the
 CC polynucleotide for a time sufficient to produce the glucan in the
 CC regenerated plant, where the vector contains a transit sequence from
 CC ribulose biphosphate carboxylase small subunit, waxy, brittle-1 and
 CC chlorophyll AB binding protein to produce a transgenic plant, and glucan
 CC is produced in the amyloplast of potato or the vacuole of sugar beet.
 CC Glucans are useful as substitutes for and additions to modified starch
 CC and latexes in paper manufacture. Unlike prior art techniques, which
 CC require input materials that produce chemical effluents, paper
 CC manufacture utilizing the glucan produced by GTF, which utilizes
 CC biologically produced input materials, is more cost-effective and
 CC environmentally friendly. Moreover, glucans also exhibit thermoplastic
 CC properties and impart gloss to the paper during coating step.
 CC The present sequence represents a GTFB mutant of the invention.
 CC Note: The present sequence is not shown in the specification but
 CC was created by the indexer using the GTFB sequence appearing as AAU98C27
 CC and the information in claim 36.
 XX
 XX
 SQ Sequence 1475 AA;
 Query Match 100.0%; Score 119; DB 23; Length 1475;
 Best Local Similarity 100.0%; Pred. No. 5,4e-09;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 ANDHLSILEAWSDNDTPYLHD 21
 DB 481 ANDHLSILEAWSDNDTPYLHD 501
 RESULT 4
 AAU98031
 ID AAU98031 standard; Protein; 1475 AA.
 XX
 AC AAU98031;
 XX
 DT 27-AUG-2002 (first entry)

XX
 DE S. mutans glucosyltransferase GTFB mutant D457N.
 XX
 KW Glucosyltransferase; GTFB; transgenic plant; paper sizing;
 KW coating composition; glucan; starch; latex; thermoplastic molecule;
 KW amyloplast; vacuole; paper manufacture; mutant; muten.
 XX
 OS Streptococcus mutans.
 OS Synthetic.
 XX
 XX Key Location/Qualifiers
 FT Misc-difference 457 /note="Wild-type Asp substituted by Asn"
 FT
 XX
 XX US2002031826-A1.
 XX
 XX 14-MAR-2002.
 XX
 XX 19-DEC-2000; 2000US-0740274.
 XX
 XX 11-DEC-1998; 98US-0210361.
 XX 07-JUN-1995; 95US-0478704.
 XX 07-JUN-1995; 95US-0482711.
 XX 07-JUN-1995; 95US-0485243.
 XX 16-JAN-1998; 98US-0007999.
 XX 16-JAN-1998; 98US-0008172.
 XX 20-JAN-1998; 98US-0009620.
 XX
 XX (NICH/) NICHOLS S E.
 XX
 XX Nichols SE;
 XX WPI; 2002-414332/44.
 XX
 DR
 XX
 PT Glucosyltransferase B or D protein useful for producing a glucan useful
 PT as substitutes for and additions to modified starch and latexes in
 PT paper manufacture, comprises mutations in specific positions
 XX
 XX
 PS Claim 36; Page -: 44pp; English.
 XX
 CC The invention an isolated protein comprising a glucosyltransferase
 CC (GTF) B polypeptide having changes at position from 1448V, D457N,
 CC D567T, K1014T, D457N/D567T, D457N/D571K, D567T/D571K,
 CC D567T/D571K/K1014T, T448V/D457N/D567T/D571K/K779Q/K1014T,
 CC Y169A/Y170A/Y171A, and K779Q or a GTF D polypeptide having
 CC changes at positions from T589D, T589E, N471D, N471D/T589D, and
 CC N471D/T589E. Also included are a glucan produced by the GTF mutant,
 CC an isolated polynucleotide which encodes P1 or P2, or its complementary
 CC polynucleotide, a ribonucleic acid sequence encoding the GTF mutant,
 CC an expression cassette comprising the polynucleotide operably linked to a
 CC promoter, a vector comprising the expression cassette, host cell
 CC introduced with the vector, a transgenic plant comprising the
 CC vector, a seed or tuber from the transgenic plant, a paper sizing and/or
 CC coating composition comprising a glucan produced in a plant transformed
 CC with a gene encoding the mutant GTF, wild type or, starch, a latex,
 CC thermoplastic molecule or their combinations or glucan and starch where
 CC the glucan is produced in the amyloplast and/or vacuole or a maize line
 CC deficient in starch biosynthesis, transformed with a gene encoding a
 CC glucosyltransferase B or D enzyme, wild-type or mutant and a paper
 CC comprising the glucan (paper sizing/coating agent). The vector is useful
 CC for producing a glucan in a plant. The method comprises transforming a
 CC plant cell with the vector, growing the plant cell under plant growing
 CC conditions to produce a regenerated plant and inducing expression of the
 CC polynucleotide for a time sufficient to produce the glucan in the
 CC regenerated plant, where the vector contains a transit sequence from
 CC ribulose biphosphate carboxylase small subunit, waxy, brittle-1 and
 CC chlorophyll AB binding protein to produce a transgenic plant, and glucan
 CC is produced in the amyloplast of potato or the vacuole of sugar beet.
 CC Glucans are useful as substitutes for and additions to modified starch
 CC and latexes in paper manufacture. Unlike prior art techniques, which
 CC require input materials that produce chemical effluents, paper
 CC manufacture utilizing the glucan produced by GTF, which utilizes
 CC biologically produced input materials, is more cost-effective and

environmentally friendly. Moreover, glucans also exhibit thermoplastic properties and impart gloss to the paper during coating step.

CC The present sequence represents a GTFB mutant of the invention.

CC Note: The present sequence is not shown in the specification but was created by the indexer using the GTFB sequence appearing as AAU98027 and the information in claim 36.

XX

Sequence 1475 AA;

Query Match 100.0%; Score 119; DB 23; Length 1475;
Best Local Similarity 100.0%; Pred. No. 5.4e-09;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 ANDHLSILEAWSNDNTPYLHD 21
481 ANDHLSILEAWSNDNTPYLHD 501

RESULT 5
AAU98032 standard; Protein; 1475 AA.

AC AAU98032;
DT 27-AUG-2002 (first entry)

S. mutans glucosyltransferase GTFB mutant D567T.

Glucosyltransferase; GTFB; transgenic plant; paper sizing;
coating composition; glucan; starch; latex; thermoplastic molecule;
amyloplast; vacuole; paper manufacture; mutant; mutein.

Streptococcus mutans.
Synthetic.

Key Location/Qualifiers
Misc-difference 567 /note= "Wild-type Asp substituted by Thr"

US2002031826-A1.

14-MAR-2002.

19-DEC-2000; 2000US-0740274.

11-DEC-1998; 98US-0210361.
07-JUN-1995; 95US-0478704.
07-JUN-1995; 95US-0482711.
07-JUN-1995; 95US-0485243.
16-JAN-1998; 98US-0007999.
16-JAN-1998; 98US-0008172.
20-JAN-1998; 98US-0009620.

(NICH/) NICHOLS S E.

Nichols SE;

WPI; 2002-414332/44.

Glucosyltransferase B or D protein useful for producing a glucan useful as substitutes for and additions to modified starch and latexes in paper manufacture, comprises mutations in specific positions -

Claim 36; Page -; 44pp; English.

The invention an isolated protein comprising a glucosyltransferase (GTF) B polypeptide having changes at position from 1448V, D457N, D567T, K1014T, D457N/D567T, D457N/D571K, D567T/D571K, D567T/D571K/K1014T, I448V/D457N/D567T/D571K/K1014T, Y169N/Y170A/Y171A, and K779Q or a GTF D polypeptide having changes at positions from T589D, T589E, N471D, N471D/T589D, and N471D/T589E. Also included are a glucan produced by the GTF mutant, an isolated polynucleotide which encodes P1 or P2, or its complementary

polynucleotide, a ribonucleic acid sequence encoding the GTF mutant, an expression cassette comprising the polynucleotide operably linked to a promoter, a vector comprising the expression cassette, host cell introduced with the vector, a transgenic plant comprising the vector, a seed or tuber from the transgenic plant, a paper sizing and/or coating composition comprising a glucan produced in a plant transformed with a gene encoding the mutant GTF, wild type or, starch, a latex, CC thermoplastic molecule or their combinations or glucan and starch where the glucan is produced in the amyloplast and/or vacuole or a maize line deficient in starch biosynthesis, transformed with a gene encoding a glucosyltransferase B or D enzyme, wild-type or mutant and a paper comprising the glucan (paper sizing/coating agent). The vector is useful for producing a glucan in a plant. The method comprises transforming a plant cell with the vector, growing the plant cell under plant growing conditions to produce a regenerated plant and inducing expression of the polynucleotide for a time sufficient to produce the glucan in the regenerated plant, where the vector contains a transit sequence from ribulose biphosphate carboxylase small subunit, waxy, brittle-1 and chlorophyll AB binding protein of potato or the vacuole of sugar beet.

CC is produced in the amyloplast of potato and additions to modified starch

CC Glucans are useful as substitutes for and additions to modified starch and latexes in paper manufacture. Unlike prior art techniques, which require input materials that produce chemical effluents, paper manufacture utilizing the glucan produced by GTF, which utilizes biologically produced input materials, is more cost-effective and environmentally friendly. Moreover, glucans also exhibit thermoplastic properties and impart gloss to the paper during coating step.

CC The present sequence represents a GTFB mutant of the invention.

CC Note: The present sequence is not shown in the specification but was created by the indexer using the GTFB sequence appearing as AAU98027 and the information in claim 36.

Sequence 1475 AA;

Query Match 100.0%; Score 119; DB 23; Length 1475;
Best Local Similarity 100.0%; Pred. No. 5.4e-09;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 ANDHLSILEAWSNDNTPYLHD 21
481 ANDHLSILEAWSNDNTPYLHD 501

RESULT 6
AAU98033 standard; Protein; 1475 AA.

AC AAU98033;
DT 27-AUG-2002 (first entry)

S. mutans glucosyltransferase GTFB mutant K1014T.

Glucosyltransferase; GTFB; transgenic plant; paper sizing;
coating composition; glucan; starch; latex; thermoplastic molecule;
amyloplast; vacuole; paper manufacture; mutant; mutein.

Streptococcus mutans.
Synthetic.

Key Location/Qualifiers
Misc-difference 1014 /note= "Wild-type Lys substituted by Thr"

US2002031826-A1.

14-MAR-2002.

19-DEC-2000; 2000US-0740274.

11-DEC-1998; 98US-0210361.
07-JUN-1995; 95US-0478704.
07-JUN-1995; 95US-0482711.

PR 07-JUN-1995; 9SUS-0485243.
PR 16-JAN-1998; 9BUS-0007999.
PR 16-JAN-1998; 9BUS-0008172.
PR 20-JAN-1998; 9BUS-0009620.

XX
XX
XX (NICH/) NICHOLS S E.
PI Nichols SE;
PI
DR WPI; 2002-414332/44.

XX Glucosyltransferase B or D protein useful for producing a glucan useful
PT as substitutes for and additions to modified starch and latexes in
PT paper manufacture, comprises mutations in specific positions -
PS Claim 36; Page -: 44pp; English.

CC The invention an isolated protein comprising a glucosyltransferase
CC (GTF) B polypeptide having changes at position from I448V, D457N,
CC D567T, K1014T, D457N/D567T, D457N/D571K, D567T/D571K,
CC D567T/D571K/K1014T, I448V/D457N/D567T/D571K/K779Q/K1014T,
CC Y169A/Y170A/Y171A, and K479Q or a GTF D polypeptide having
CC changes at positions from T589D, T589E, N471D, N471D/T589D, and
CC N471D/T589E. Also included are a glucan produced by the GTF mutant,
CC an isolated polynucleotide which encodes P1 or P2, or its complementary
CC polynucleotide, a ribonucleic acid sequence encoding the GTF mutant,
CC an expression cassette comprising the polynucleotide operably linked to a
CC promoter, a vector comprising the expression cassette, host cell
CC introduced with the vector, a transgenic plant comprising the
CC vector, a seed or tuber from the transgenic plant, a paper string and/or
CC coating composition comprising a glucan produced in a plant transformed
CC with a gene encoding the mutant GTF, wild type or, starch, a latex,
CC thermoplastic molecule or their combinations or glucan and starch where
CC the glucan is produced in the amyloplast and/or vacuole or a maize line
CC deficient in starch biosynthesis, transformed with a gene encoding a
CC glucosyltransferase B or D enzyme, wild-type or mutant and a paper
CC comprising the glucan (paper sizing/coating agent). The vector is useful
CC for producing a glucan in a plant. The method comprises transforming a
CC plant cell with the vector, growing the plant cell under plant growing
CC conditions to produce a regenerated plant and inducing expression of the
CC polynucleotide for a time sufficient to produce the glucan in the
CC regenerated plant, where the vector contains a transit sequence from
CC ribulose biphosphate carboxylase small subunit, waxy, brittle-1 and
CC chlorophyll AB binding protein to produce a transgenic plant, and glucan
CC is produced in the amyloplast of potato or the vacuole of sugar beet.
CC Glucans are useful as substitutes for and additions to modified starch
CC and latexes in paper manufacture. Unlike prior art techniques, which
CC require input materials that produce chemical effluents, paper
CC manufacture utilizing the glucan produced by GTF, which utilizes
CC biologically produced input materials, is more cost-effective and
CC environmentally friendly. Moreover, glucans also exhibit thermoplastic
CC properties and impart gloss to the paper during coating step.
CC The present sequence represents a GTFB mutant of the invention.
CC Note: The present sequence is not shown in the specification but
CC was created by the index using the GTFB sequence appearing as AAU98027
CC and the information in claim 36.
CC XX

SQ Sequence 1475 AA;

Query Match 100.0%; Score 119; DB 23; Length 1475;
Best Local Similarity 100.0%; Pred. No. 5.4e-09;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Dd 481 ANDHSLLEAWSDNDTPYLHD 501
1 ANDHSLLEAWSDNDTPYLHD 21

RESULT 7
ID AAU98034 standard; Protein; 1475 AA.
AC AAU98034;

XX	27-AUG-2002	(first entry)
DT	S. mutans glucosyltransferase GTFB mutant D457N/D567T.	
XX		
DE		
XX	Glucosyltransferase; GTFB; transgenic plant; paper sizing;	
KW	coating composition; glucan; starch; latex; thermoplastic molecule;	
KW	amyloplast; vacuole; paper manufacture; mutant; mutain.	
XX		
OS	Streptococcus mutans.	
OS	Synthetic.	
XX		
FH	Key	
FT	Misc-difference 457	
FT	/note= "Wild-type Asp substituted by Asn"	
FT	Misc-difference 567	
FT	/note= "Wild-type Asp substituted by Thr"	
XX		
FN	US2002031826-A1.	
XX		
PD	14-MAR-2002.	
XX		
PF	19-DEC-2000; 2000US-0740274.	
XX		
PR	11-DEC-1998; 98US-0210361.	
PR	07-JUN-1995; 95US-0478704.	
PR	07-JUN-1995; 95US-0482711.	
PR	07-JUN-1995; 95US-0485243.	
PR	16-JAN-1998; 98US-0007999.	
PR	16-JAN-1998; 98US-0008172.	
PR	20-JAN-1998; 98US-0009620.	
XX		
PA	(NICH/) NICHOLS S E.	
XX		
PI	Nichols SE;	
DR	WPI; 2002-414332/44.	
XX		
PT	Glucosyltransferase B or D protein useful for producing a glucan useful	
PT	as substitutes for and additions to modified starch and latexes in	
PT	paper manufacture, comprises mutations in specific positions -	
XX		
PS	Claim 36; Page -; 44pp; English.	
XX		
CC	The invention an isolated protein comprising a glucosyltransferase	
CC	(GTF) B polypeptide having changes at position from I448V, D457N,	
CC	D567T, K1014T, D457N/D567T, D457N/D571K, D567T/D571K,	
CC	D567T/D571K/K1014T, I448V/D457N/D567T/D571K/K779Q/K1014T,	
CC	I459A/Y170A/Y171A, and K779Q or a GTF D polypeptide having	
CC	changes at positions from T589D, T589E, M471D, N471D/T589D, and	
CC	M471D/T589E. Also included are a glucan produced by the GTF mutant,	
CC	an isolated polynucleotide which encodes PI or P2, or its complementary	
CC	polynucleotide, a ribonucleic acid sequence encoding the GTF mutant,	
CC	an expression cassette comprising the expression cassette, host cell	
CC	promoter, a vector comprising the expression cassette, host cell	
CC	introduced with the vector, a transgenic plant comprising the	
CC	vector, a seed or tuber from the transgenic plant, a paper sizing and/or	
CC	coating composition comprising a glucan produced in a plant transformed	
CC	with a gene encoding the mutant GTF, wild type or, starch, a latex,	
CC	thermoplastic molecule or their combinations or glucan and starch where	
CC	the glucan is produced in the amyloplast and/or vacuole or a maize line	
CC	deficient in starch biosynthesis, transformed with a gene encoding a	
CC	glucosyltransferase B or D enzyme, wild-type or mutant and a paper	
CC	comprising the glucan (paper sizing/coating agent). The vector is useful	
CC	for producing a glucan in a plant. The method comprises transforming a	
CC	plant cell with the vector, growing the plant cell under plant growing	
CC	conditions to produce a regenerated plant and inducing expression of the	
CC	polynucleotide for a time sufficient to produce the glucan in the	
CC	regenerated plant, where the vector contains a transit sequence from	
CC	ribulose biphosphate carboxylase small subunit, waxy, brittle-1 and	
CC	chlorophyll AB binding protein to produce a transgenic plant, and glucan	
CC	is produced in the amyloplast of potato or the vacuole of sugar beet.	
CC	Glucans are useful as substitutes for and additions to modified starch	

CC and latexes in paper manufacture. Unlike prior art techniques, which
 CC require input materials that produce chemical effluents, paper
 CC manufacture utilizing the glucan produced by GTF, which utilizes
 CC biologically produced input materials, is more cost-effective and
 CC environmentally friendly. Moreover, glucans also exhibit thermoplastic
 CC properties and impart gloss to the paper during coating step.
 CC The present sequence represents a GTFB mutant of the invention.
 CC Note: The present sequence is not shown in the specification but
 CC was created by the indexer using the GTFB sequence appearing as AAU98027
 CC and the information in claim 36.

CC Sequence 1475 AA;

Query Match 100.0%; Score 119; DB 23; Length 1475;
 Best Local Similarity 100.0%; Pred. No. 5,4e-09;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ANDHSLTEAMSDNDPTPLHD 21
 DB 481 ANDHSLTEAMSDNDPTPLHD 501

RESULT 8
 AAU98035 ID AAU98035 standard; Protein; 1475 AA.

XX AC AAU98035;

XX DT 27-AUG-2002 (first entry)

XX DE S. mutans glucosyltransferase GTFB mutant D457N/D571K.

XX KM Glucosyltransferase; GTFB; transgenic plant; paper sizing;
 XX coating composition; glucan; starch; latex; thermoplastic molecule;

XX KW amyloplast; vacuole; paper manufacture; mutant; mutain.

XX OS Streptococcus mutans.

XX OS Synthetic.

XX FH Key Location/Qualifiers

XX FT Misc-difference 457 /note= "Wild-type Asp substituted by Asn"

XX FT Misc-difference 571 /note= "Wild-type Asp substituted by Lys"

XX FT US2002031826-A1.

XX PD 14-MAR-2002.

XX PF 19-DEC-2000; 2000US-0740274.

XX PR 11-DEC-1998; 98US-0210361.

XX PR 07-JUN-1995; 95US-0478704.

XX PR 07-JUN-1995; 95US-0482711.

XX PR 07-JUN-1995; 95US-0485243.

XX PR 16-JAN-1998; 98US-0007999.

XX PR 16-JAN-1998; 98US-0008172.

XX PR 20-JAN-1998; 98US-0009620.

XX PA (NICH/) NICHOLS S. E.

XX PI Nichols SE;

XX DR WPI; 2002-414332/44.

XX XX Glucosyltransferase B or D protein useful for producing a glucan useful
 XX PT as substitutes for and additions to modified starch and latexes in
 XX PT paper manufacture, comprises mutations in specific positions -

XX RS Claim 36; Page -; 44pp; English.

XX XX The invention an isolated protein comprising a glucosyltransferase
 CC (GTF) B polypeptide having changes at position from I448V, D457N,

CC D567T, K1014T, D457N/D567T, D457N/D571K, D567T/D571K,
 CC D567T/D571K/K1014T, 1448V/D457N/D567T/D571K/K7790/K1014T,
 CC Y169A/Y170A/Y171A, and K7790 or a GTF D polypeptide having
 CC changes at positions from T589D, T589E, N471D, N471D/T589D
 CC N471D/T589E. Also included are a glucan produced by the GTF mutant,
 CC an isolated polynucleotide which encodes P1 or P2, or its complementary
 CC polynucleotide, a ribonucleic acid sequence encoding the GTF mutant,
 CC an expression cassette comprising the polynucleotide operably linked to a
 CC promoter, a vector comprising the expression cassette, host cell
 CC introduced with the vector, a transgenic plant comprising the
 CC vector, a seed or tuber from the transgenic plant, a paper sizing and/or
 CC coating composition comprising a glucan produced in a plant transformed
 CC with a gene encoding the mutant GTF, wild type or, starch, a latex,
 CC thermoplastic molecule in the amyloplast and/or vacuole or a maize line
 CC deficient in starch biosynthesis, transformed with a gene encoding a
 CC glucosyltransferase B or D enzyme, wild-type or mutant and a paper
 CC comprising the glucan (paper sizing/coating agent). The vector is useful
 CC for producing a glucan in a plant. The method comprises transforming a
 CC plant cell with the vector, growing the plant cell under plant growing
 CC conditions to produce a regenerative plant and inducing expression of the
 CC polynucleotide for a time sufficient to produce the glucan in the
 CC regenerated plant, where the vector contains a transit sequence from
 CC ribulose biphosphate carboxylase small subunit, waxy, brittle-1 and
 CC chlorophyll AB binding protein to produce a transgenic plant, and glucan
 CC is produced in the amyloplast of potato or the vacuole of sugar beet.
 CC Glucans are useful as substitutes for and additions to modified starch
 CC and latexes in paper manufacture. Unlike prior art techniques, which
 CC require input materials that produce chemical effluents, paper
 CC manufacture utilizing the glucan produced by GTF, which utilizes
 CC biologically produced input materials, is more cost-effective and
 CC environmentally friendly. Moreover, glucans also exhibit thermoplastic
 CC properties and impart gloss to the paper during coating step.
 CC The present sequence represents a GTFB mutant of the invention.
 CC Note: The present sequence is not shown in the specification but
 CC was created by the indexer using the GTFB sequence appearing as AAU98027
 CC and the information in claim 36.

XX SQ Sequence 1475 AA;

Query Match 100.0%; Score 119; DB 23; Length 1475;
 Best Local Similarity 100.0%; Pred. No. 5,4e-09;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ANDHSLTEAMSDNDPTPLHD 21
 DB 481 ANDHSLTEAMSDNDPTPLHD 501

RESULT 9

AAU98036 ID AAU98036 standard; Protein; 1475 AA.

XX AC AAU98036;

XX DT 27-AUG-2002 (first entry)

XX DE S. mutans glucosyltransferase GTFB mutant D567T/D571K.

XX KM Glucosyltransferase; GTFB; transgenic plant; paper sizing;
 XX coating composition; glucan; starch; latex; thermoplastic molecule;

XX KW amyloplast; vacuole; paper manufacture; mutant; mutain.

XX OS Streptococcus mutans.

XX OS Synthetic.

XX FH Key Location/Qualifiers

XX FT Misc-difference 567 /note= "Wild-type Asp substituted by Thr"

XX FT Misc-difference 571 /note= "Wild-type Asp substituted by Lys"

XX FT US2002031826-A1.

14-MAR-2002.
19-DEC-2000; 2000US-0740274.
11-DEC-1998; 98US-0210361.
07-JUN-1995; 95US-0478704.
07-JUN-1995; 95US-0482711.
07-JUN-1995; 95US-0485243.
16-JAN-1998; 98US-0007999.
16-JAN-1998; 98US-0008172.
20-JAN-1998; 98US-0009620.
(NICH/) NICHOLS S E.
Nichols SE;
WPI; 2002-414332/44.
Glucosyltransferase B or D protein useful for producing a glucan useful as substitutes for and additions to modified starch and latexes in paper manufacture, comprises mutations in specific positions -
Claim 36; Page -; 44pp; English.
The invention an isolated protein comprising a glucosyltransferase (GTF) B polypeptide having changes at position from 1448V, D457N, D567T, K1014T, D457N/D567T, D457N/D571K, D567T/D571K, D567T/D571K/K1014T, I448V/D457N/D567T/D571K/K1014T, Y169A/Y170A/Y171A, and K779Q or a GTF D polypeptide having changes at positions from T589D, T589E, N471D, N471D/T589D, and N471D/T589E. Also included are a glucan produced by the GTF mutant, an isolated polynucleotide which encodes P1 or P2, or its complementary polynucleotide, a ribonucleic acid sequence encoding the GTF mutant, an expression cassette comprising the polynucleotide operably linked to a promoter, a vector comprising the expression cassette, host cell introduced with the vector, a transgenic plant comprising the vector, a seed or tuber from the transgenic plant, a paper sizing and/or coating composition comprising a glucan produced in a plant transformed with a gene encoding the mutant GTF, wild type or starch, a latex, thermoplastic molecule or their combinations or glucan and starch where the glucan is produced in the amyloplast and/or vacuole of a maize line deficient in starch biosynthesis, transformed with a gene encoding a glucosyltransferase B or D enzyme, wild-type or mutant and a paper comprising the glucan (paper sizing/coating agent). The vector is useful for producing a glucan in a plant. The method comprises transforming a plant cell with the vector, growing the plant cell under plant growing conditions to produce a regenerated plant and inducing expression of the polynucleotide for a time sufficient to produce the glucan in the regenerated plant, where the vector contains a transit sequence from ribulose biphosphate carboxylase small subunit, waxy, brittle-1 and chlorophyll AB binding protein to produce a transgenic plant, and glucan is produced in the amyloplast of potato or the vacuole of sugar beet. Glucans are useful as substitutes for and additions to modified starch and latexes in paper manufacture. Unlike prior art techniques, which require input materials that produce chemical effluents, paper manufacture utilizing the glucan produced by GTF, which utilizes biologically produced input materials, is more cost-effective and environmentally friendly. Moreover, glucans also exhibit thermoplastic properties and impart gloss to the paper during coating step. The present sequence represents a GTF mutant of the invention. Note: The present sequence is not shown in the specification but was created by the indexer using the GTFB sequence appearing as AAU98027 and the information in claim 36.

Sequence 1475 AA;
Query Match 100.0%; Score 119; DB 23; Length 1475;
Best Local Similarity 100.0%; Pred. No. 5.4e-09;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1 ANDHLSILEAWSNDNTPYLHD 21
|||||

DB 481 ANDHLSILEAWSNDNTPYLHD 501
RESULT 10
AAU98037
ID AAU98037 standard; Protein; 1475 AA.
XX
AC AAU98037;
XX
DT 27-AUG-2002 (first entry)
XX
DE S. mutans glucosyltransferase GTFB mutant D567T/D571K/K1014T.
XX
KW Glucosyltransferase; GTFB; transgenic plant; paper sizing;
XX coating composition; glucan; starch; latex; thermoplastic molecule;
XX amyloplast; vacuole; paper manufacture; mutant; mutcin.
OS Streptococcus mutans.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Misc-difference 567 /note= "Wild-type Asp substituted by Thr"
FT Misc-difference 571 /note= "Wild-type Asp substituted by Lys"
FT Misc-difference 1014 /note= "Wild-type Lys substituted by Thr"
XX
PN US2002031826-A1.
XX
PD 14-MAR-2002.
XX
PF 19-DEC-2000; 2000US-0740274.
XX
PR 11-DEC-1998; 98US-0210361.
XX 07-JUN-1995; 95US-0478704.
XX 07-JUN-1995; 95US-0482711.
XX 07-JUN-1995; 95US-0485243.
XX 16-JAN-1998; 98US-0007999.
XX 16-JAN-1998; 98US-0008172.
XX 20-JAN-1998; 98US-0009620.
XX
PA (NICH/) NICHOLS S E.
XX
PI Nichols SE;
XX
DR WPI; 2002-414332/44.
XX
PT Glucosyltransferase B or D protein useful for producing a glucan useful as substitutes for and additions to modified starch and latexes in paper manufacture, comprises mutations in specific positions -
XX
PS Claim 36; Page -; 44pp; English.
XX
XX The invention an isolated protein comprising a glucosyltransferase (GTF) B polypeptide having changes at position from 1448V, D457N, D567T, K1014T, D457N/D567T, D457N/D571K, D567T/D571K, D567T/D571K/K1014T, I448V/D457N/D567T/D571K/K1014T, Y169A/Y170A/Y171A, and K779Q or a GTF D polypeptide having changes at positions from T589D, T589E, N471D, N471D/T589D, and N471D/T589E. Also included are a glucan produced by the GTF mutant, an isolated polynucleotide which encodes P1 or P2, or its complementary polynucleotide, a ribonucleic acid sequence encoding the GTF mutant, an expression cassette comprising the polynucleotide operably linked to a promoter, a vector comprising the expression cassette, host cell introduced with the vector, a transgenic plant comprising the vector, a seed or tuber from the transgenic plant, a paper sizing and/or coating composition comprising a glucan produced in a plant transformed with a gene encoding the mutant GTF, wild type or starch, a latex, thermoplastic molecule or their combinations or glucan and starch where the glucan is produced in the amyloplast and/or vacuole of a maize line deficient in starch biosynthesis, transformed with a gene encoding a glucosyltransferase B or D enzyme, wild-type or mutant and a paper

CC comprising the glucan (paper sizing/coating agent). The vector is useful
 CC for producing a glucan in a plant. The method comprises transforming a
 CC plant cell with the vector, growing the plant cell under plant growing
 CC conditions to produce a regenerated plant and inducing expression of the
 CC polynucleotide for a time sufficient to produce the glucan in the
 CC regenerated plant, where the vector contains a transit sequence from
 CC ribulose biphosphate carboxylase small subunit, waxy, brittle-1 and
 CC chlorophyll AB binding protein to produce a transgenic plant, and glucan
 CC is produced in the amyloplast of potato or the vacuole of sugar beet.
 CC Glucans are useful as substitutes for and additions to modified starch
 CC and latexes in paper manufacture. Unlike prior art techniques, which
 CC require input materials that produce chemical effluents, paper
 CC manufacture utilizing the glucan produced by GFP, which utilizes
 CC biologically produced input materials, is more cost-effective and
 CC environmentally friendly. Moreover, glucans also exhibit thermoplastic
 CC properties and impart gloss to the paper during coating step.
 CC The present sequence represents a GFPB mutant of the invention.
 CC Note: The present sequence is not shown in the specification but
 CC was created by the indexer using the GFPB sequence appearing as AAU98027
 CC and the information in claim 36.

XX
 CC Sequence 1475 AA;
 SQ

Query Match 100.0%; Score 119; DB 23; Length 1475;
 Best Local Similarity 100.0%; Pred. No. 5.4e-09;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ANDHSLTEAWSDNDTPYLHD 21
 |||||
 Db 481 ANDHSLTEAWSDNDTPYLHD 501

RESULT 11
 AAU98038
 ID AAU98038 standard; Protein; 1475 AA.
 AC AAU98038;
 XX
 XX
 DT 27-AUG-2002 (first entry)
 DE S. mutans GFPB mutant 1448V/D457N/D567T/D571K/K779Q/K1014T.
 XX
 XX Glucosyltransferase; GFPB; transgenic plant; paper sizing;
 KM coating composition; glucan; starch; latex; thermoplastic molecule;
 KM amyloplast; vacuole; paper manufacture; mutant; mutlein.
 XX
 OS Streptococcus mutans.
 OS Synthetic.
 XX
 XX Key Location/Qualifiers
 FH Misc-difference 448 /note= "Wild-type Ile substituted by Val"
 FT Misc-difference 457 /note= "Wild-type Asp substituted by Asn"
 FT Misc-difference 567 /note= "Wild-type Asp substituted by Thr"
 FT Misc-difference 571 /note= "Wild-type Asp substituted by Lys"
 FT Misc-difference 779 /note= "Wild-type Lys substituted by Gln"
 FT Misc-difference 1014 /note= "Wild-type Lys substituted by Thr"
 FT
 XX
 XX PN US2002031826-A1.
 XX
 XX PD 14-MAR-2002.
 XX
 XX PF 19-DEC-2000; 2000US-0740274.
 XX
 XX PR 11-DEC-1998; 98US-0210361.
 PR 07-JUN-1995; 95US-0478704.
 PR 07-JUN-1995; 95US-0482711.
 PR 07-JUN-1995; 95US-0485243.

PR 16-JAN-1998; 98US-0007999.
 PR 16-JAN-1998; 98US-0008172.
 PR 20-JAN-1998; 98US-0009620.
 XX
 XX
 XX (NICH/) NICHOLS S E.
 XX
 XX Nichols SE;
 DR WPI; 2002-414332/44.
 XX
 XX
 PT Glucosyltransferase B or D protein useful for producing a glucan useful
 PT as substitutes for and additions to modified starch and latexes in
 PT paper manufacture, comprises mutations in specific positions
 PS
 XX Claim 36; Page -; 44pp; English.
 XX
 XX The invention an isolated protein comprising a glucosyltransferase
 CC (GTF) B polypeptide having changes at position from 1448V, D457N,
 CC D567T, K1014T, D457N/D567T, D457N/D571K, D567T/D571K,
 CC D567T/D571K/K1014T, 1448V/D457N/D567T/D571K/K779Q/K1014T,
 CC Y169A/Y170A/Y171A, and K779Q or a GFP D polypeptide having
 CC changes at positions from T589D, T589E, N471D, N471D/T589D, and
 CC N471D/T589E. Also included are a glucan produced by the GFP mutant,
 CC an isolated polynucleotide which encodes P1 or P2, or its complementary
 CC polynucleotide, a ribonucleic acid sequence encoding the GFP mutant, a
 CC an expression cassette comprising the polynucleotide operably linked to a
 CC promoter, a vector comprising the expression cassette, host cell
 CC introduced with the vector, a transgenic plant comprising the
 CC vector, a seed or tuber from the transgenic plant, a paper sizing and/or
 CC coating composition comprising a glucan produced in a plant transformed
 CC with a gene encoding the mutant GFP, wild type or, starch, a latex,
 CC thermoplastic molecule or their combinations or glucan and starch where
 CC the glucan is produced in the amyloplast and/or vacuole or a maize line
 CC deficient in starch biosynthesis, transformed with a gene encoding a
 CC glucosyltransferase B or D enzyme, wild-type or mutant and a paper
 CC comprising the glucan (paper sizing/coating agent). The vector is useful
 CC for producing a glucan in a plant. The method comprises transforming a
 CC plant cell with the vector, growing the plant cell under plant growing
 CC conditions to produce a regenerated plant and inducing expression of the
 CC polynucleotide for a time sufficient to produce the glucan in the
 CC regenerated plant, where the vector contains a transit sequence from
 CC ribulose biphosphate carboxylase small subunit, waxy, brittle-1 and
 CC chlorophyll AB binding protein to produce a transgenic plant, and glucan
 CC is produced in the amyloplast of potato or the vacuole of sugar beet.
 CC Glucans are useful as substitutes for and additions to modified starch
 CC and latexes in paper manufacture. Unlike prior art techniques, which
 CC require input materials that produce chemical effluents, paper
 CC manufacture utilizing the glucan produced by GFP, which utilizes
 CC biologically produced input materials, is more cost-effective and
 CC environmentally friendly. Moreover, glucans also exhibit thermoplastic
 CC properties and impart gloss to the paper during coating step.
 CC The present sequence represents a GFPB mutant of the invention.
 CC Note: The present sequence is not shown in the specification but
 CC was created by the indexer using the GFPB sequence appearing as AAU98027
 CC and the information in claim 36.

XX
 XX Sequence 1475 AA;
 SQ

Query Match 100.0%; Score 119; DB 23; Length 1475;
 Best Local Similarity 100.0%; Pred. No. 5.4e-09;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ANDHSLTEAWSDNDTPYLHD 21
 |||||
 Db 481 ANDHSLTEAWSDNDTPYLHD 501

RESULT 12
 AAU98039
 ID AAU98039 standard; Protein; 1475 AA.
 AC AAU98039;
 XX
 XX
 XX

27-AUG-2002 (first entry)
 S. mutans glucosyltransferase GTFB mutant YYY169-171AAA.
 Glucosyltransferase; GTFB; transgenic plant; paper sizing;
 coating composition; glucan; starch; latex; thermoplastic molecule;
 amyloplast; vacuole; paper manufacture; mutant; mutein.
 Streptococcus mutans.
 Synthetic.
 Key Location/Qualifiers
 Misc-difference 169..171
 /note="Wild-type Tyr-Tyr-Tyr substituted by
 Ala-Ala-Ala"
 US2002031826-A1.
 14-MAR-2002.
 19-DEC-2000; 2000US-0740274.
 11-DEC-1998; 98US-0210361.
 07-JUN-1995; 95US-0478704.
 07-JUN-1995; 95US-0482711.
 07-JUN-1995; 95US-0485243.
 16-JAN-1998; 98US-0007999.
 16-JAN-1998; 98US-0008172.
 20-JAN-1998; 98US-0009620.
 (NICH/) NICHOLS S E.
 Nichols SE;
 WPI; 2002-414332/44.
 Glucosyltransferase B or D protein useful for producing a glucan useful
 as substitutes for and additions to modified starch and latexes in
 paper manufacture, comprises mutations in specific positions -
 Claim 36; Page -; 44pp; English.
 The invention an isolated protein comprising a glucosyltransferase
 (GTF) B polypeptide having changes at position from 1448V, D457N,
 D567T, K1014T, D457N/D567T, D457N/D571K, D567T/D571K,
 D567T/D571K/K1014T, T448V/D457N/D567T/D571K/K779Q/K1014T,
 Y169A/Y170A/Y171A, and K779Q or a GTF D polypeptide having
 changes at positions from T589D, T589E, N471D, N471D/T589D, and
 N471D/T589E. Also included are a glucan produced by the GTF mutant,
 an isolated polynucleotide which encodes P1 or P2, or its complementary
 polynucleotide, a ribonucleic acid sequence encoding the GTF mutant,
 an expression cassette comprising the polynucleotide operably linked to a
 promoter, a vector comprising the expression cassette, host cell
 introduced with the vector, a transgenic plant comprising the
 vector, a seed or tuber from the transgenic plant, a paper sizing and/or
 coating composition comprising a glucan produced in a plant transformed
 with a gene encoding the mutant GTF, wild type or, starch, a latex,
 thermoplastic molecule or their combinations or glucan and starch where
 the glucan is produced in the amyloplast and/or vacuole or a maize line
 deficient in starch biosynthesis, transformed with a gene encoding a
 glucosyltransferase B or D enzyme, wild-type or mutant and a paper
 comprising the glucan (paper sizing/coating agent). The vector is useful
 for producing a glucan in a plant. The method comprises transforming a
 plant cell with the vector, growing the plant cell under plant growing
 conditions to produce a regenerated plant and inducing expression of the
 polynucleotide for a time sufficient to produce the glucan in the
 regenerated plant, where the vector contains a transit sequence from
 ribulose biphosphate carboxylase small subunit, waxy, brittle-1 and
 chlorophyll AB binding protein to produce a transgenic plant, and glucan
 is produced in the amyloplast of potato or the vacuole of sugar beet.
 Glucans are useful as substitutes for and additions to modified starch
 and latexes in paper manufacture. Unlike prior art techniques, which
 require input materials that produce chemical effluents, paper

Manufacture utilizing the glucan produced by GTF, which utilizes
 biologically produced input materials, is more cost-effective and
 environmentally friendly. Moreover, glucans also exhibit thermoplastic
 properties and impart gloss to the paper during coating step.
 The present sequence represents a GTFB mutant of the invention.
 Note: The present sequence is not shown in the specification but
 was created by the indexer using the GTFB sequence appearing as AAU98027
 and the information in claim 36.
 Sequence 1475 AA;
 Query Match 100.0%; Score 119; DB 23; Length 1475;
 Best Local Similarity 100.0%; Pred. No. 5.4e-09;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 ANDHLSILEAMSDNDPTYLHD 21
 DB 481 ANDHLSILEAMSDNDPTYLHD 501
 RESULT 13
 AAU98040
 ID AAU98040 standard; Protein; 1475 AA.
 AC AAU98040;
 DT 27-AUG-2002 (first entry)
 DE S. mutans glucosyltransferase GTFB mutant K779Q.
 KW Glucosyltransferase; GTFB; transgenic plant; paper sizing;
 coating composition; glucan; starch; latex; thermoplastic molecule;
 amyloplast; vacuole; paper manufacture; mutant; mutein.
 Streptococcus mutans.
 Synthetic.
 Key Location/Qualifiers
 Misc-difference 779
 /note="Wild-type Lys substituted by Gln"
 US2002031826-A1.
 14-MAR-2002.
 19-DEC-2000; 2000US-0740274.
 11-DEC-1998; 98US-0210361.
 07-JUN-1995; 95US-0478704.
 07-JUN-1995; 95US-0482711.
 07-JUN-1995; 95US-0485243.
 16-JAN-1998; 98US-0007999.
 16-JAN-1998; 98US-0008172.
 20-JAN-1998; 98US-0009620.
 (NICH/) NICHOLS S E.
 Nichols SE;
 WPI; 2002-414332/44.
 Glucosyltransferase B or D protein useful for producing a glucan useful
 as substitutes for and additions to modified starch and latexes in
 paper manufacture, comprises mutations in specific positions -
 Claim 36; Page -; 44pp; English.
 The invention an isolated protein comprising a glucosyltransferase
 (GTF) B polypeptide having changes at position from 1448V, D457N,
 D567T, K1014T, D457N/D567T, D457N/D571K, D567T/D571K,
 D567T/D571K/K1014T, T448V/D457N/D567T/D571K/K779Q/K1014T,
 Y169A/Y170A/Y171A, and K779Q or a GTF D polypeptide having
 changes at positions from T589D, T589E, N471D, N471D/T589D, and

CC N471D/T589E. Also included are a glucan produced by the GTF mutant,
 CC an isolated polynucleotide which encodes P1 or P2, or its complementary
 CC polynucleotide, a ribonucleic acid sequence encoding the GTF mutant,
 CC an expression cassette comprising the polynucleotide operably linked to a
 CC promoter, a vector comprising the expression cassette, host cell
 CC introduced with the vector, a transgenic plant comprising the
 CC vector, a seed or tuber from the transgenic plant, a paper sizing and/or
 CC coating composition comprising a glucan produced in a plant transformed
 CC with a gene encoding the mutant GTF, wild type or, starch, a latex,
 CC thermoplastic molecule or their combinations and/or vacuole or a maize line
 CC the glucan is produced in the amyloplast and/or vacuole of a maize line
 CC deficient in starch biosynthesis, transformed with a gene encoding a
 CC glucosyltransferase B or D enzyme, wild-type or mutant and a paper
 CC comprising the glucan (paper sizing/coating agent). The vector is useful
 CC for producing a glucan in a plant. The method comprises transforming a
 CC plant cell with the vector, growing the plant cell under plant growing
 CC conditions to produce a regenerated plant and inducing expression of the
 CC polynucleotide for a time sufficient to produce the glucan in the
 CC regenerated plant, where the vector contains a transit sequence from
 CC ribulose biphosphate carboxylase small subunit, waxy, brittle-1 and
 CC chlorophyll AB binding protein to produce a transgenic plant, and glucan
 CC is produced in the amyloplast of potato or the vacuole of sugar beet.
 CC Glucans are useful as substitutes for and additions to modified starch
 CC and latexes in paper manufacture. Unlike prior art techniques, which
 CC require input materials that produce chemical effluents, paper
 CC manufacture utilizing the glucan produced by GTF, which utilizes
 CC biologically produced input materials, is more cost-effective and
 CC environmentally friendly. Moreover, glucans also exhibit thermoplastic
 CC properties and impart gloss to the paper during coating step.
 CC The present sequence represents a GTFB mutant of the invention.
 CC Note: The present sequence is not shown in the specification but
 CC was created by the indexer using the GTFB sequence appearing as AAU98027
 CC and the information in claim 36.

CC Sequence 1475 AA;

SO Query Match 100.0%; Score 119; DB 23; Length 1475;
 Best Local Similarity 100.0%; Pred. No. 5.4e-09;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ANDHSLILEAWSNDPTYLHD 21
 DB 481 ANDHSLILEAWSNDPTYLHD 501

RESULT 14
 AAU79284
 ID AAU79284 standard; Protein; 1476 AA.

AC AAU79284;
 XX
 AC
 XX
 DT 13-AUG-2002 (first entry)
 XX
 DE Streptococcus mutans monoclonal antibody-related protein #1.
 XX
 KW Antibody; dental caries; water insoluble glucan synthetase;
 KW anti-carries; glucosyl transferase-B; immunotherapy.
 XX
 OS Streptococcus mutans.
 XX
 JP2002114709-A.
 PN
 XX
 PD 16-APR-2002.
 XX
 PF 04-OCT-2000; 2000UP-0304889.
 XX
 PR 04-OCT-2000; 2000UP-0304889.
 XX
 PA (UYNI-) UNIV NIPPON.
 XX
 DR WPI; 2002-448101/48.
 XX
 PT Anti-carries agent composed of a monoclonal antibody against an

PT inhibitory enzyme against water insoluble glucan synthetase of glucosyl
 PT transferase-B (GTF-B) of Streptococcus mutans
 XX
 PS Claim 3; Page 13-16; 28pp; Japanese.

CC The invention relates to a monoclonal antibody against dental caries and
 CC an anti-carries agent composed of a monoclonal antibody produced by
 CC Streptococcus mutans, particularly mouse-hybridoma MHP126 (FERM P-17566)
 CC or mouse-hybridoma MHP136 (FERM P-17567), against an enzyme having
 CC inhibitive activity against water insoluble glucan synthetase of glucosyl
 CC transferase-B. The monoclonal antibody specifically inhibits water
 CC insoluble glucan synthetase of Streptococcus mutans produced glucosyl
 CC transferase-B and is used in the immunotherapy of dental caries. This
 CC sequence represents a Streptococcus mutans monoclonal antibody-related
 CC protein.

CC Sequence 1476 AA;

SO Query Match 100.0%; Score 119; DB 23; Length 1476;
 Best Local Similarity 100.0%; Pred. No. 5.4e-09;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ANDHSLILEAWSNDPTYLHD 21
 DB 481 ANDHSLILEAWSNDPTYLHD 501

RESULT 15
 AAU98028
 ID AAU98028 standard; Protein; 1375 AA.

XX AAU98028;

DT 27-AUG-2002 (first entry)

DE S. mutans glucosyltransferase GTFB.

XX Glucosyltransferase; GTFB; transgenic plant; paper sizing;
 KW coating composition; glucan; starch; latex; thermoplastic molecule;
 KW amyloplast; vacuole; paper manufacture.

XX Streptococcus mutans.

OS US2002031826-A1.

PN 14-MAR-2002.

PD 19-DEC-2000; 2000US-0740274.

PF 11-DEC-1998; 98US-0210361.
 PR 07-JUN-1995; 95US-0478704.
 PR 07-JUN-1995; 95US-0482711.
 PR 16-JAN-1998; 98US-0007999.
 PR 16-JAN-1998; 98US-0008172.
 PR 20-JAN-1998; 98US-0009620.

XX (NICH/) NICHOLS S E.

PA Nichols SE;

XX WPI; 2002-414332/44.

DR N-PSDB; ABR52939.

XX Glucosyltransferase B or D protein useful for producing a glucan useful
 PT as substitutes for and additions to modified starch and latexes in
 PT paper manufacture, comprises mutations in specific positions

XX Disclosure; Page 30-33; 44pp; English.

CC The invention an isolated protein comprising a glucosyltransferase
 CC (GTF) B polypeptide having changes at position from I448V, D457N,
 CC D567T, K1014T, D457N/D567T, D457N/D571K, D567T/D571K,

CC D567T/D571K/K1014T, 1448V/D457N/D567T/D571K/K779Q/K1014T,
 CC Y169A/Y170A/Y171A, and K779Q or a GFP D polypeptide having
 CC changes at positions from T589D, T589E, N471D, N471D/T589D, and
 CC N471D/T589E. Also included are a glucan produced by the GFP mutant,
 CC an isolated polynucleotide which encodes P1 or P2, or its complementary
 CC polynucleotide, a ribonucleic acid sequence encoding the GFP mutant,
 CC an expression cassette comprising the polynucleotide operably linked to a
 CC promoter, a vector comprising the expression cassette, host cell
 CC introduced with the vector, a transgenic plant comprising the
 CC vector, a seed or tuber from the transgenic plant, a paper sizing and/or
 CC coating composition comprising a glucan produced in a plant transformed
 CC with a gene encoding the mutant GFP, wild type or, starch, a latex,
 CC thermoplastic molecule or their combinations or glucan and starch where
 CC the glucan is produced in the amyloplast and/or vacuole or a maize line
 CC deficient in starch biosynthesis, transformed with a gene encoding a
 CC glucosyltransferase B or D enzyme, wild-type or mutant and a paper
 CC comprising the glucan (paper sizing/coating agent). The vector is useful
 CC for producing a glucan in a plant. The method comprises transforming a
 CC plant cell with the vector, growing the plant cell under plant growing
 CC conditions to produce a regenerated plant and inducing expression of the
 CC polynucleotide for a time sufficient to produce the glucan in the
 CC regenerated plant, where the vector contains a transit sequence from
 CC ribulose biphosphate carboxylase small subunit, waxy, brittle-1 and
 CC chlorophyll AB binding protein to produce a transgenic plant, and glucan
 CC is produced in the amyloplast of potato or the vacuole of sugar beet.
 CC Glucans are useful as substitutes for and additions to modified starch
 CC and latexes in paper manufacture. Unlike prior art techniques, which
 CC require input materials that produce chemical effluents, paper
 CC manufacture utilizing the glucan produced by GFP, which utilizes
 CC biologically produced input materials, is more cost-effective and
 CC environmentally friendly. Moreover, glucans also exhibit thermoplastic
 CC properties and impart gloss to the paper during coating step.
 CC The present sequence represents GTRC.

XX Sequence 1375 AA;

Query Match 92.4%; Score 110; DB 23; Length 1375;

Best Local Similarity 95.2%; Pred. No. 1.1e-07;

Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ANDHSLTEAWSDNDPTYLHD 21
 |||||
 Db 507 ANDHSLTEAWSYNDPTYLHD 527

Search completed: November 13, 2003, 09:38:26
 Job time : 38.5166 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 13, 2003, 09:31:40 ; Search time 13.0379 Seconds
(without alignments)
154.898 Million cell updates/sec

Title: US-09-290-049A-10
Perfect score: 119
Sequence: 1 ANDHSLTEAMSDNDPTPLHD 21

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues
Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	119	100.0	1475	2	B33135	gtfB protein precursor
2	110	92.4	1375	2	JT0345	dextranucrase (EC
3	108	90.8	1592	2	A38175	glucosyltransferase
4	78	65.5	1365	2	A41483	glucosyltransferase
5	75	63.0	1431	2	A45866	dextranucrase (EC
6	74	62.2	1508	2	T31098	probable dextran
7	66	55.5	1449	2	T30857	glucosyltransferase
8	66	55.5	1449	2	T30552	glucosyltransferase
9	63	52.9	1577	2	T30858	glucosyltransferase
10	58	48.7	1518	2	A44811	glyceroldehyde-3-P
11	57	47.9	331	2	B48445	glucosyltransferase
12	57	47.9	1599	2	S22737	glucosyltransferase
13	52	43.7	4848	2	T30289	pristinamycin I sy
14	50	42.0	347	2	T48610	hypothetical prote
15	49.5	41.6	549	2	T03983	hypothetical prote
16	49	41.2	302	2	AG0597	conserved hypotret
17	49	41.2	336	1	OCBE40	BGLF2 protein - hu
18	49	41.2	601	2	B87028	pyruvate, phosphat
19	48	40.3	490	2	H70538	probable ppk prot
20	47.5	39.9	331	1	DEUTGC	glyceroldehyde-3-P
21	47.5	39.9	519	2	C86372	hypothetical prote
22	47	39.5	175	2	C86205	hypothetical prote
23	47	39.5	378	2	T04254	hypothetical prote
24	47	39.5	524	2	D82220	conserved hypotret
25	47	39.5	525	2	T40088	RhoGEF domain cont
26	46.5	39.1	418	2	D90506	4-aminobutyrate am
27	46	38.7	215	2	S55978	hypothetical prote
28	46	38.7	335	2	E86568	glyceroldehyde-3-P
29	46	38.7	335	2	B72053	glyceroldehyde-3-P

30	46	38.7	400	2	E98138	hypothetical prote
31	46	38.7	417	2	A83149	glucosyltransferase
32	46	38.7	476	1	S58229	salicylate biosynt
33	45.5	38.2	96	2	B81786	conserved hypotret
34	45	37.8	148	2	B95908	hypothetical prote
35	45	37.8	236	2	S67732	guanidinocacetate N
36	45	37.8	309	2	D95905	conserved hypotret
37	45	37.8	947	2	B86362	hypothetical prote
38	45	37.8	1090	2	AG1749	glycosidase homolo
39	45	37.8	1091	2	AF1380	glycosidase homolo
40	45	37.8	1159	2	T43461	probable phosphodi
41	44.5	37.4	210	2	H83332	conserved hypotret
42	44	37.0	155	2	AD2392	transcription regu
43	44	37.0	247	1	P01778	glyceroldehyde-3-P
44	44	37.0	331	1	DEECG3	glyceroldehyde-3-P
45	44	37.0	331	2	H90939	glyceroldehyde-3-P

ALIGNMENTS

RESULT 1

B33135 gtfB protein precursor - Streptococcus mutans

C/Species: Streptococcus mutans
C/Date: 23-Oct-1990 #sequence_revision 23-Oct-1990 #text_change 15-Oct-1999
C/Accession: B33135; A33128
R/Shiroza, T.; Ueda, S.; Kuramitsu, H.K.
J. Bacteriol. 169, 4263-4270, 1987

A/Title: Sequence analysis of the gtfB gene from Streptococcus mutans.

A/Reference number: A33135; MID:87308013; PMID:3040685

A/Accession: B33135

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-1475 <SH1>

A/Cross-references: GB:M17361; NID:G153639; PIDN:AAA88588.1; PID:G153640

R/Shiroza, T.; Ueda, S.; Kuramitsu, H.K.
submitted to the Protein Sequence Database, September 1990

A/Reference number: A33128

A/Accession: A33128

A/Status: preliminary; not compared with conceptual translation

A/Molecule type: DNA

A/Residues: 1-171,173-641,'N',643-1475 <SH2>

A/Experimental source: strain GS-5

C/Superfamily: cpl repeat homology

F:1096-1115/Domain: cpl repeat homology <CP1>

F:1224-1243/Domain: cpl repeat homology <CP2>

F:1289-1308/Domain: cpl repeat homology <CP3>

F:1354-1373/Domain: cpl repeat homology <CP4>

F:1419-1438/Domain: cpl repeat homology <CP5>

Query Match

Best Local Similarity 100.0%; Score 119; DB 2; Length 1475;

Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ANDHSLTEAMSDNDPTPLHD 21

Db 481 ANDHSLTEAMSDNDPTPLHD 501

RESULT 2

JT0345 dextranucrase (EC 2.4.1.5) precursor - Streptococcus mutans (strain GS-5)

N/Alternate names: sucrose 6-glucosyltransferase

C/Species: Streptococcus mutans

C/Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 24-Sep-1999

C/Accession: JT0345; C33135

R/Ueda, S.; Shiroza, T.; Kuramitsu, H.K.
Gene 69, 101-109, 1988

A/Title: Sequence analysis of the gtfC gene from Streptococcus mutans GS-5.

A/Reference number: JT0345; MID:89137980; PMID:2976010

A/Accession: JT0345

A/Molecule type: DNA

A;Residues: 1-1375 <UED>
 A;Experimental source: GS-5
 R;Shiroza, T.; Ueda, S.; Kuramitsu, H.K.
 U. Bacteriol. 169, 4263-4270, 1987
 A;Title: Sequence analysis of the *gtfB* gene from *Streptococcus mutans*.
 A;Reference number: A31135, MUID:87308013, PMID:3040685
 A;Accession: C31135
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-349 <SHI>
 A;Cross-references: GB:M17361
 A;Genetics:
 C;Function:
 A;Description: catalyzes the synthesis of both water-soluble and water-insoluble glucans
 C;Superfamily: cpl repeat homology
 C;Keywords: duplication; glycosyltransferase; hexosyltransferase
 F;1-34/Domain: signal sequence #status predicted <SIG>
 F;35-1375/Product: glycosyltransferase #status predicted <MAT>
 F;1126-1145/Domain: cpl repeat homology <CP1>
 F;1253-1272/Domain: cpl repeat homology <CP2>
 F;1318-1337/Domain: cpl repeat homology <CP3>

Query Match 92.4%; Score 110; DB 2; Length 1375;
 Best Local Similarity 95.2%; Pred. No. 1.2e-08;
 Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ANDHSLILEAMSDNDTPYLHD 21
 Db 507 ANDHSLILEAMSDNDTPYLHD 527

RESULT 3

A38175
 glucosyltransferase precursor - *Streptococcus sobrinus*

C;Species: *Streptococcus sobrinus*
 C;Date: 28-Aug-1992 #sequence_revision 28-Aug-1992 #text_change 15-Oct-1995
 C;Accession: A38175
 R;Abo, H.; Mesumura, T.; Kodama, T.; Ohta, H.; Fukui, K.; Kato, K.; Kagawa, H.
 U. Bacteriol. 173, 989-996, 1991
 A;Title: Peptide sequences for sucrose splitting and glucan binding within *Streptococcus*
 A;Reference number: A38175; MUID:91123227; PMID:1704006
 A;Accession: A38175
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-1592 <ABO>
 A;Cross-references: GB:D90213; NID:9217032; PIDN:BA14241.1; PID:41014946; PID:9217033
 C;Superfamily: cpl repeat homology
 F;1093-1112/Domain: cpl repeat homology <CP1>
 F;1222-1241/Domain: cpl repeat homology <CP2>
 F;1287-1306/Domain: cpl repeat homology <CP3>
 F;1330-1351/Domain: cpl repeat homology <CP4>
 F;1352-1371/Domain: cpl repeat homology <CP5>
 F;1402-1420/Domain: cpl repeat homology <CP6>
 F;1465-1484/Domain: cpl repeat homology <CP7>
 F;1513-1532/Domain: cpl repeat homology <CP8>

Query Match 90.8%; Score 108; DB 2; Length 1592;
 Best Local Similarity 85.7%; Pred. No. 3e-08;
 Matches 18; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 ANDHSLILEAMSDNDTPYLHD 21
 Db 477 ANNHVSIVEAMSDNDTPYLHD 497

RESULT 4

A41483
 glucosyltransferase (EC 2.4.1.-) *gtfs* precursor - *Streptococcus sobrinus*
 C;Species: *Streptococcus sobrinus*
 C;Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 15-Oct-1999
 C;Accession: A41483
 R;Gillmore, K.S.; Russell, R.R.B.; Ferretti, J.J.

Infect. Immun. 58, 2452-2458, 1990

A;Title: Analysis of the *Streptococcus downei* *gtfs* gene, which specifies a glucosyltrans
 A;Reference number: A41483; MUID:90316655; PMID:2142479
 A;Accession: A41483
 A;Molecule type: DNA
 A;Residues: 1-1365 <GLU>
 A;Cross-references: GB:M30943; NID:9153652; PIDN:AAA26898.1; PID:9153653
 C;Genetics:
 A;Gene: *gtfs*
 C;Superfamily: cpl repeat homology
 C;Keywords: glycosyltransferase; hexosyltransferase

Query Match 65.5%; Score 78; DB 2; Length 1365;
 Best Local Similarity 71.4%; Pred. No. 0.0012;
 Matches 15; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 ANDHSLILEAMSDNDTPYLHD 21
 Db 467 AIDHSLILEAMSDNDTPYLHD 487

RESULT 5

A45866
 dextranucrase (EC 2.4.1.5) precursor - *Streptococcus mutans*

C;Species: *Streptococcus mutans*
 C;Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 07-May-1999
 C;Accession: A45866
 R;Honda, O.; Kato, C.; Kuramitsu, H.K.
 U. Gen. Microbiol. 136, 2099-2105, 1990
 A;Title: Nucleotide sequence of the *Streptococcus mutans* *gtfd* gene encoding the glucosyl
 A;Reference number: A45866; MUID:91100958; PMID:2148600
 A;Accession: A45866
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-1431 <HON>
 A;Cross-references: GB:M29296
 C;Superfamily: cpl repeat homology
 C;Keywords: glycosyltransferase; hexosyltransferase
 F;181-201/Domain: cpl repeat homology <CP1>
 F;1127-1146/Domain: cpl repeat homology <CP2>
 F;1192-1211/Domain: cpl repeat homology <CP3>
 F;1257-1276/Domain: cpl repeat homology <CP4>
 F;1277-1297/Domain: cpl repeat homology <CP5>
 F;1321-1340/Domain: cpl repeat homology <CP6>
 F;1341-1361/Domain: cpl repeat homology <CP7>
 F;1385-1404/Domain: cpl repeat homology <CP8>

Query Match 63.0%; Score 75; DB 2; Length 1431;

Best Local Similarity 71.4%; Pred. No. 0.0036;
 Matches 15; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 ANDHSLILEAMSDNDTPYLHD 21
 Db 495 AINHLILEAMSDNDTPYLHD 515

RESULT 6

T31098
 probable dextranucrase (EC 2.4.1.5), extracellular - *Leuconostoc mesenteroides*

C;Species: *Leuconostoc mesenteroides*
 C;Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 11-May-2000
 C;Accession: T31098
 R;Monchois, V.; Remaud-Simeon, M.; Monsan, P.; Willemot, R.M.
 FEBS Microbiol. Lett. 159, 307-315, 1998
 A;Title: Cloning and sequencing of a gene coding for an extracellular dextranucrase (D
 A;Reference number: Z20981; MUID:98164374; PMID:9503626
 A;Accession: T31098
 A;Status: preliminary; translated from GB/EMBL/DBD
 A;Molecule type: DNA
 A;Residues: 1-1508 <MON>
 A;Cross-references: EMBL:AF030129; NID:92766611; PID:92766612; PIDN:AAB95453.1
 A;Experimental source: strain NRRL B-1299
 C;Genetics:

A:Gene: darB
C:Function:
A:Description: produces dextran composed only of alpha(1-6) glucosidic bonds
C:Keywords: glycosyltransferase; hexosyltransferase

Query Match 62.2%; Score 74; DB 2; Length 1508;
Best Local Similarity 66.7%; Pred. No. 0.0055;
Matches 14; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 1 ANDHSLTLEAWSNDTPYLHD 21
DB 563 ANQHSLTLEAWSNDPEYKCD 583

RESULT 7
T30857
glucosyltransferase - Streptococcus salivarius

C:Species: Streptococcus salivarius
C:Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 22-Oct-1999
C:Accession: T30857
R:Simpson, C.L.; Giffard, P.M.; Jacques, N.A.
Infect. Immun. 63, 609-621, 1995

A:Title: Streptococcus salivarius ATCC 25975 possesses at least two genes coding for pri
A:Reference number: Z20909; MUID:95122197; PMID:7822030
A:Accession: T30857
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1449 <SIM>
A:Cross-references: EMBL:L35495; NID:g662378; PID:g662379; PIDN:AAC41412.1
C:Keywords:
A:Gene: gtfI

Query Match 55.5%; Score 66; DB 2; Length 1449;
Best Local Similarity 72.2%; Pred. No. 0.092;
Matches 13; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 4 HSLTLEAWSNDTPYLHD 21
DB 539 HSLTLEAWSNDAYNED 556

RESULT 8
T30552

glucosyltransferase N - Streptococcus salivarius (fragment)
C:Species: Streptococcus salivarius
C:Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 22-Oct-1999
C:Accession: T30552
R:Jaife, R.I.

A:Submitted to the EMBL Data Library, February 1998
A:Description: Streptococcus salivarius VI477 gtfN.
A:Reference number: Z20854
A:Accession: T30552
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1449 <GAF>
A:Cross-references: EMBL:AF049609; NID:g2935545; PID:g2935546; PIDN:AAC05156.1
C:Keywords:
A:Gene: gtfN

Query Match 55.5%; Score 66; DB 2; Length 1449;
Best Local Similarity 72.2%; Pred. No. 0.092;
Matches 13; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 4 HSLTLEAWSNDTPYLHD 21
DB 539 HSLTLEAWSNDAYNED 556

RESULT 9
T30858

glucosyltransferase - Streptococcus salivarius
C:Species: Streptococcus salivarius
C:Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 22-Oct-1999

C:Accession: T30858
R:Simpson, C.L.; Giffard, P.M.; Jacques, N.A.
Infect. Immun. 63, 609-621, 1995

A:Title: Streptococcus salivarius ATCC 25975 possesses at least two genes coding for pri
A:Reference number: Z20909; MUID:95122197; PMID:7822030
A:Accession: T30858
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1577 <SIM>
A:Cross-references: EMBL:L35928; NID:g662380; PID:g662381; PIDN:AAC41413.1
C:Keywords:
A:Gene: gtfm

Query Match 52.9%; Score 63; DB 2; Length 1577;
Best Local Similarity 72.2%; Pred. No. 0.3;
Matches 13; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 4 HSLTLEAWSNDTPYLHD 21
DB 591 HSLTLEAWSNDHOYKCD 608

RESULT 10
A44811

glucosyltransferase (EC 2.4.1.-) I - Streptococcus salivarius
C:Species: Streptococcus salivarius
C:Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 15-Oct-1999
C:Accession: A44811; S22726; S28809
R:Giffard, P.M.; Simpson, C.L.; Milward, C.F.; Jacques, N.A.
J. Gen. Microbiol. 137, 2577-2593, 1991
A:Title: Molecular characterization of a cluster of at least two glucosyltransferase gen
A:Reference number: A44811; MUID:92148377; PMID:1838391
A:Accession: A44811
A:Molecule type: DNA
A:Residues: 1-1518 <GIF>

A:Cross-references: EMBL:Z11873; NID:947526; PIDN:CAA77900.1; PID:947527
A:Note: sequence extracted from NCBI backbone (NCBIN:81050, NCBI:P:81052)
C:Keywords: glycosyltransferase; hexosyltransferase
C:Superfamily: cpl repeat homology
A:Gene: gtfU
C:Superfamily: cpl repeat homology
F:1307-1326/Domain: cpl repeat homology <CP4>

Query Match 48.7%; Score 58; DB 2; Length 1518;
Best Local Similarity 66.7%; Pred. No. 1.7;
Matches 10; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 4 HSLTLEAWSNDTPYLHD 18
DB 504 HSLTLEAWSNDNHX 518

RESULT 11
B48445

glyceraldehyde-3-phosphate dehydrogenase (phosphorylating) (EC 1.2.1.12) - Leishmania me
C:Species: Leishmania mexicana
C:Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 03-Jun-2002
C:Accession: B48445; S25142
R:Hanaert, V.; Blaauw, M.; Kohl, L.; Allert, S.; Oppendes, F.R.; Michels, P.A.M.
Mol. Biochem. Parasitol. 55, 115-126, 1992
A:Title: Molecular analysis of the cytosolic and glycosomal glyceraldehyde-3-phosphate de
A:Reference number: A48445; MUID:93063042; PMID:1435864
A:Accession: B48445
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-331 <HAN>
A:Cross-references: EMBL:X65220; NID:g95552; PIDN:CAA46323.1; PID:g95553
C:Superfamily: glyceraldehyde-3-phosphate dehydrogenase
C:Keywords: oxidoreductase

Query Match 47.9%; Score 57; DB 2; Length 331;
Best Local Similarity 52.6%; Pred. No. 0.4;
Matches 10; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using SW model

Run on: November 13, 2003, 09:31:40 ; Search time 7.56398 Seconds
(without alignments)
130.561 Million cell updates/sec

Title: US-09-290-049A-10

Perfect score: 119
Sequence: 1 ANDHUS1EFAWSDNDTPYLDH 21

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	119	100.0	1476	1 GTFB_STRMU	P08987 streptococc
2	110	92.4	1455	1 GTFB_STRMU	P13470 streptococc
3	108	90.8	1592	1 GTF2_STRMO	P27470 streptococc
4	108	90.8	1597	1 GTF1_STRMO	P11001 streptococc
5	78	65.5	1365	1 GTF5_STRMO	P29336 streptococc
6	75	63.0	1462	1 GTFD_STRMU	P49331 streptococc
7	57	47.9	330	1 G3PC_LEIME	Q01558 leishmania
8	49	41.2	302	1 YBHK_SALTI	P58587 salmoneila
9	49	41.2	302	1 YBHK_SALTY	P58586 salmoneila
10	49	41.2	336	1 U1A6_EBV	P10321 epstein-bar
11	47.5	39.9	330	1 G3PC_TKRYB	P10097 trypanosoma
12	47	39.5	337	1 G3P_COCCO	Q081h3 coccidioid
13	46	38.7	275	1 SC35_YEAST	P53271 saccharomyc
14	46	38.7	335	1 G3P_CHLPN	Q92770 chlamydia p
15	46	38.7	476	1 PCHN_PSEAR	Q51508 pseudomonas
16	45	37.8	236	1 GANT_HUMAN	Q14353 homo sapien
17	45	37.8	385	1 GNTF_FUSOX	P46239 fusarium ox
18	44.5	37.4	470	1 AMP2_STRCO	Q060394 streptomyce
19	44	37.0	234	1 G3P_MELGA	Q57672 melaleucis g
20	44	37.0	330	1 G3P1_ECOLI	P06597 escherichia
21	44	37.0	330	1 G3P1_SALTY	P24165 salmoneila
22	44	37.0	332	1 G3P_CHICK	P00335 gallus galli
23	44	37.0	332	1 G3P_COULI	O57479 columba liv
24	44	37.0	337	1 G3PC_MAIZE	P08735 zea mays (m
25	44	37.0	337	1 G3PD_MAIZE	Q09054 zea mays (m
26	44	37.0	337	1 G3PD_HOYU	P26517 hordeum vul
27	44	37.0	337	1 G3P_HOYU	P19089 cryphonectr
28	44	37.0	363	1 G3P1_VACOR	P80534 jaculus ori
29	44	37.0	816	1 SUS2_WAIZE	P49336 zea mays (m
30	43.5	36.6	115	1 TKNI_RABIT	P41540 oryctolagus
31	43.5	36.6	130	1 TKNI_MESAT	Q60541 mesocricetu
32	43.5	36.6	130	1 TKNI_MOUSE	P41539 mus muscula
33	43.5	36.6	130	1 TKNI_RAT	P06767 rattus norv

34	43.5	36.6	491	1 G6PD_ERWCH	P37986 erwina chr
35	43	36.1	174	1 PADC_VIBCH	Q9Kpx2 vibrio chol
36	43	36.1	235	1 GANT_RAT	P10868 rattus norv
37	43	36.1	236	1 GANT_MOUSE	Q35969 bos taurus
38	43	36.1	320	1 G3P_BOVIN	P10096 bos taurus
39	43	36.1	322	1 G3P_SHREP	Q28554 ovis aries
40	43	36.1	332	1 G3P_PIG	P00355 sus scrofa
41	43	36.1	332	1 G3P_RABIT	P46406 oryctolagus
42	43	36.1	334	1 G3P2_HUMAN	P04406 homo sapien
43	43	36.1	337	1 G3P_AJECU	Q9htx1 ajellomyces
44	43	36.1	337	1 G3P_MYCCE	P47543 mycoplasma
45	43	36.1	424	1 IVD_RAT	P12007 rattus norv

ALIGNMENTS

RESULT 1
GTFB_STRMU STANDARD: PRT; 1476 AA.
ID GTFB_STRMU
AC P08987; 069381; 069384; 069387; 069390; 069396;
DT 01-NOV-1988 (Rel. 09, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Glucosyltransferase-I precursor (EC 2.4.1.5) (GTF-I) (Dextranucrase)
GN (Sucrose 6-glucosyltransferase).
OS Streptococcus mutans.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1309;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=GS-5.
RX MEDLINE=87308013; PubMed=3040685;
RA Shiroza T., Ueda S., Kuramitsu H.K.;
RT "Sequence analysis of the gtfB gene from Streptococcus mutans."
RL J. Bacteriol. 169:4263-4270(1987).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=MT4239 / Serotype C, MT4245 / Serotype E, MT4251 / Serotype F,
RX MT4467 / Serotype B, and MT8148 / Serotype G;
RA MEDLINE=98231643; PubMed=9570124;
RA Fujiwara T., Terao Y., Hoshino T., Kawabata S., Ooshima T., Sobue S.,
RA Kimura S., Hamada S.;
RT "Molecular analyses of glucosyltransferase genes among strains of
Streptococcus mutans."
RL FEWS Microbiol. Lett. 161:331-336(1998).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=UA159 / ATCC 700610 / Serotype C;
RX MEDLINE=22295063; PubMed=12397186;
RA Aylton D., McNamee W.M., McLaughlin R.E., Savic G., Chang J.,
RA Carlson W.B., Primeaux C., Tian R., Kenton S., Jia H., Lin S., Qian Y.,
RA Li S., Zhu H., Najjar F., Lai H., White U., Roe B.A., Fretwell J.J.;
RT "Genome sequence of Streptococcus mutans UA159, a cariogenic dental
pathogen."
RL Proc. Natl. Acad. Sci. U.S.A. 99:14434-14439(2002).
CC - FUNCTION: PRODUCTION OF EXTRACELLULAR GLUCANS, THAT ARE THOUGHT
TO PLAY A KEY ROLE IN THE DEVELOPMENT OF THE DENTAL PLAQUE BECAUSE
OF THEIR ABILITY TO ADHERE TO SMOOTH SURFACES AND MEDIANE THE
AGGREGATION OF BACTERIAL CELLS AND FOOD DEBRIS.
CC - CATALYTIC ACTIVITY: Sucrose + ((1,6)-alpha-D-glucosyl) (N) = D-
fructose + ((1,6)-alpha-D-glucosyl) (N+1).
CC - SUBCELLULAR LOCATION: Secreted.
CC - DISEASE: DENTAL CARIES.
CC - MISCELLANEOUS: GTF-I SYNTHESIZES WATER-INSOLUBLE GLUCANS (ALPHA
1,3-LINKED GLUCOSE AND SOME 1,6 LINKAGES), GTF-S SYNTHESIZES
WATER-SOLUBLE GLUCANS (ALPHA 1,6-GLUCOSE). GTF-SI SYNTHESIZES BOTH
FORMS OF GLUCANS.
CC - SIMILARITY: TO OTHER GLUCOSYLTRANSFERASES AND SOME TO A GLUCAN-
BINDING PROTEIN FROM S. MUTANS.
CC - SIMILARITY: Contains 10 cell wall binding repeats.


```

CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.1sb-sib.ch/announce/
CC      or send an email to license@1sb-sib.ch).
CC      -----
DR      EMBL; M17361; AAA8858.1; -
DR      EMBL; D88651; BAA2610.1; -
DR      EMBL; D88654; BAA2610.1; -
DR      EMBL; D88657; BAA2610.1; -
DR      EMBL; D88660; BAA2611.1; -
DR      EMBL; D89977; BAA2611.1; -
DR      EMBL; AE014940; AAN58705.1; -
DR      InterPro; IPR002479; CW_binding.
DR      InterPro; IPR003118; Glyco_hydro_70.
DR      Pfam; PF01473; CW_binding_1; 1.
DR      Pfam; PF02324; Glyco_hydro_70; 1.
DR      Transferrase; Glycosyltransferase; Signal; Repeat; Dental carries;
KM      Complete proteome.
FT      SIGNAL 1 34
FT      CHAIN 35 1476
FT      DOMAIN 35 1051
FT      DOMAIN 1097 1476
FT      REPEAT 1097 1130
FT      REPEAT 1161 1470
FT      REPEAT 1161 1210
FT      REPEAT 1225 1275
FT      REPEAT 1290 1340
FT      REPEAT 1355 1405
FT      REPEAT 1420 1470
FT      REPEAT 62 62
FT      REPEAT 65 65
FT      REPEAT 68 68
FT      VARIANT 78 78
FT      VARIANT 86 86
FT      VARIANT 89 89
FT      VARIANT 168 168
FT      VARIANT 276 276
FT      VARIANT 399 399
FT      VARIANT 474 474
FT      VARIANT 512 512
FT      VARIANT 519 519
FT      VARIANT 701 701
FT      VARIANT 708 708
FT      VARIANT 938 938
FT      VARIANT 952 957
FT      VARIANT 963 964
FT      VARIANT 968 970
FT      VARIANT 1086 1086
FT      VARIANT 1158 1158
FT      VARIANT 1163 1163
FT      VARIANT 1168 1168
FT      VARIANT 1182 1182
FT      VARIANT 1234 1234
FT      VARIANT 1263 1263
FT      VARIANT 1263 1263
FT      VARIANT 1264 1264
FT      VARIANT 1272 1272
FT      VARIANT 1329 1329
FT      VARIANT 1394 1394
FT      VARIANT 1402 1402

```

```

FT      VARIANT 1459 1459
FT      VARIANT 570 570
FT      CONFLICT 800 817
FT      CONFLICT 1310 1310
FT      CONFLICT 1476 AA; 165846 MM; 9C6B09F71B4CBFC CRC64;
SQ      SEQUENCE 1476 AA; 165846 MM; 9C6B09F71B4CBFC CRC64;
Query Match 100.0%; Score 119; DB 1; Length 1476;
Best Local Similarity 100.0%; Pred. No. 1.2e-10;
Matches 21; Conservative 0; Mismatches 0; Gaps 0;
Ox 1 ANDHSTLEAWSDNDPYLHD 21
Db 481 ANDHSTLEAWSDNDPYLHD 501
RESULT 2
GTFPc STRMU STANDARD; PRT; 1455 AA.
AC P13470; O69382; O69385; O69388; O69391; O69397; F05427;
DT 01-NOV-1988 (Rel. 09, Created)
DT 28-FEB-2003 (Rel. 41, last sequence update)
DT 15-SEP-2003 (Rel. 42, last annotation update)
DE Glucosyltransferase-SI precursor (EC 2.4.1.5) (GTF-SI)
DE (Dextranucrase) (Sucrose 6-glucosyltransferase).
GN GTFP OR SMO.1005.
OS Streptococcus mutans.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1309;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=GS-5;
RX MEDLINE=89137980; PubMed=2976010;
RA Ueda S., Shiroza T., Kuramitsu H.K.;
RT "Sequence analysis of the gtfC gene from Streptococcus mutans GS-5.";
RL Gene 69:101-109(1988).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=MT4239 / Serotype C; MT4245 / Serotype E; MT4251 / Serotype F;
RX MT4467 / Serotype E; and MT8148 / Serotype C;
RX MEDLINE=98231643; PubMed=9570124;
RA Fujiwara T., Terao Y., Hoshino T., Kawabata S., Ooshima T., Sobue S.,
RA Kimura S., Hamada S.;
RT "Molecular analyses of glucosyltransferase genes among strains of
RT Streptococcus mutans.";
RL FEMS Microbiol. Lett. 161:331-336(1998).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=UAI59 / ATCC 700610 / Serotype C;
RX MEDLINE=22295063; PubMed=12397186;
RA Ajdic D., McShan W.M., McLaughlin R.E., Savic G., Chang J.,
RA Carson M.B., Primeaux C., Tian R., Kenton S., Jia H., Lin S., Qian Y.,
RA Li S., Zhu H., Najjar F., Lai H., White J., Roe B.A., Ferretti J.J.;
RT "Genome sequence of Streptococcus mutans UAI59, a cariogenic dental
RT pathogen.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:14434-14439(2002).
RN [4]
RP SEQUENCE OF 1-349 FROM N.A.
RC STRAIN=GS-5;
RX MEDLINE=87308013; PubMed=3040685;
RA Shiroza T., Ueda S., Kuramitsu H.K.;
RT "Sequence analysis of the gtfB gene from Streptococcus mutans.";
RL J. Bacteriol. 169:4263-4270(1987).
CC -1- FUNCTION: PRODUCTION OF EXTRACELLULAR GLUCANS, THAT ARE THOUGHT
CC OF THEIR ABILITY TO ADHERE TO SMOOTH SURFACES AND MEDATE THE
CC AGGREGATION OF BACTERIAL CELLS AND FOOD DEBRIS.
CC -1- CATALYTIC ACTIVITY: Sucrose + (1,6)-alpha-D-glucosyl(1) (N) = D-
CC fructose + (1,6)-alpha-D-glucosyl(1) (N) + D-
CC -1- SUBCELLULAR LOCATION: secreted.
CC -1- DISEASE: DENTAL CARRIES.

```


FT SIGNAL 1 38 POTENTIAL.
 FT CHAIN 39 1592 GLUCOSYLTRANSFERASE-I.
 FT DOMAIN 39 1044 CATALYTIC (APPROXIMATE).
 FT DOMAIN 1093 1592 GLUCAN-BINDING (APPROXIMATE).
 FT DOMAIN 1093 1592 6.5 X TANDEM REPEATS.
 FT REPEAT 1093 1142 1.
 FT REPEAT 1158 1207 2.
 FT REPEAT 1158 1207 2.
 FT REPEAT 1222 1272 3.
 FT REPEAT 1287 1337 4.
 FT REPEAT 1402 1451 5.
 FT REPEAT 1514 1563 6.
 FT REPEAT 1577 1592 7 (INCOMPLETE).
 SQ SEQUENCE 1592 AA; 176167 MW; BCOA6D079351BCF CRC64;
 Query Match 90.8%; Score 108; DB 1; Length 1592;
 Best Local Similarity 85.7%; Pred. No. 7.6e-09;
 Matches 18; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 ANDHSLTEAMSNDTPYLHD 21
 DB 477 ANNHVSIVEAMSDNTPYLHD 497

RESULT 4
 GTF1_STRDO STANDARD; PRT; 1597 AA.
 ID GTF1_STRDO
 AC P11001;
 DT 01-JUL-1989 (Rel. 11, Created)
 DT 01-JUL-1989 (Rel. 11, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Glucosyltransferase-I precursor (EC 2.4.1.5) (GTF-I) (Dextranucrase)
 DE (Sucrose 6-glucosyltransferase).
 GN GTFI.
 OS Streptococcus downei (Streptococcus sobrinus).
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 OC NCBI_TaxID=1317;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MEF28;
 RX MEDLINE=87308014; PubMed=3040686;
 RA Ferretti J.J., Galpin M.L., Russell R.R.B.;
 RT "Nucleotide sequence of a glucosyltransferase gene from Streptococcus
 RT sobrinus MEF28."
 RL J. Bacteriol. 169:4271-4278(1987).
 CC -1- FUNCTION: PRODUCTION OF EXTRACELLULAR GLUCANS, THAT ARE THOUGHT
 CC TO PLAY A KEY ROLE IN THE DEVELOPMENT OF THE DENTAL PLAQUE BECAUSE
 CC OF THEIR ABILITY TO ADHERE TO SMOOTH SURFACES AND MEDATE THE
 CC AGGREGATION OF BACTERIAL CELLS AND FOOD DEBRIS.
 CC -1- CATALYTIC ACTIVITY: Sucrose + (1,6)-alpha-D-glucosyl (N) = D-
 CC fructose + (1,6)-alpha-D-glucosyl (N+1).
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- DISEASE: DENTAL CARIES.
 CC -1- MISCELLANEOUS: GTF-I SYNTHESIZES WATER-INSOLUBLE GLUCANS (ALPHA
 CC 1,3-LINKED GLUCOSE AND SOME 1,6 LINKAGES). GTF-SI SYNTHESIZES BOTH
 CC WATER-SOLUBLE GLUCANS (ALPHA 1,6-GLUCOSE). GTF-SI SYNTHESIZES BOTH
 CC FORMS OF GLUCANS.
 CC -1- SIMILARITY: TO OTHER GLUCOSYLTRANSFERASES AND SOME TO A GLUCAN-
 CC BINDING PROTEIN FROM S. MUTANS.
 CC -1- SIMILARITY: Contains 19 cell wall binding repeats.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC EMBL; M17391; AAC63063.1; -
 CC InterPro; IPR002479; CW binding.
 CC InterPro; IPR003318; Glyco Hydro_70.
 CC Pfam; PF01473; CW_binding_1; 16.

DR Pfam; PF02324; Glyco Hydro_70; 1.
 KW Glucosyltransferase; Glycosyltransferase; Signal; Repeat; Dental caries.
 FT SIGNAL 1 38 POTENTIAL.
 FT CHAIN 39 1597 GLUCOSYLTRANSFERASE-I.
 FT DOMAIN 39 1050 CATALYTIC (APPROXIMATE).
 FT DOMAIN 1099 1597 GLUCAN-BINDING (APPROXIMATE).
 FT DOMAIN 1099 1597 1.25 A, 2 B, AND 5 AC REPEATS.
 FT REPEAT 1099 1132 A REPEAT.
 FT REPEAT 1163 1213 AC REPEAT.
 FT REPEAT 1227 1272 AC REPEAT.
 FT REPEAT 1292 1342 AC REPEAT.
 FT REPEAT 1352 1399 B REPEAT.
 FT REPEAT 1406 1455 B REPEAT.
 FT REPEAT 1465 1512 B REPEAT.
 FT REPEAT 1519 1568 AC REPEAT.
 FT REPEAT 1582 1597 A REPEAT (INCOMPLETE).
 SQ SEQUENCE 1597 AA; 177080 MW; B9586A200868798E CRC64;
 Query Match 90.8%; Score 108; DB 1; Length 1597;
 Best Local Similarity 85.7%; Pred. No. 7.6e-09;
 Matches 18; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 ANDHSLTEAMSNDTPYLHD 21
 DB 483 ANNHVSIVEAMSDNTPYLHD 503

RESULT 5
 GTF5_STRDO STANDARD; PRT; 1365 AA.
 ID GTF5_STRDO
 AC P29336;
 DT 01-DEC-1992 (Rel. 24, Created)
 DT 01-DEC-1992 (Rel. 24, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Glucosyltransferase-S precursor (EC 2.4.1.5) (GTF-S) (Dextranucrase)
 DE (Sucrose 6-glucosyltransferase).
 GN GTF5.
 OS Streptococcus downei (Streptococcus sobrinus).
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 OC NCBI_TaxID=1317;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MEF28;
 RX MEDLINE=90316655; PubMed=2142479;
 RA Gilmore K.S., Russell R.R., Ferretti J.J.;
 RT "Analysis of the Streptococcus downei gtf5 gene, which specifies a
 RT glucosyltransferase that synthesizes soluble glucans."
 RL Infect. Immun. 58:2452-2458(1990).
 CC -1- FUNCTION: PRODUCTION OF EXTRACELLULAR GLUCANS, THAT ARE THOUGHT
 CC TO PLAY A KEY ROLE IN THE DEVELOPMENT OF THE DENTAL PLAQUE BECAUSE
 CC OF THEIR ABILITY TO ADHERE TO SMOOTH SURFACES AND MEDATE THE
 CC AGGREGATION OF BACTERIAL CELLS AND FOOD DEBRIS.
 CC -1- CATALYTIC ACTIVITY: Sucrose + (1,6)-alpha-D-glucosyl (N) = D-
 CC fructose + (1,6)-alpha-D-glucosyl (N+1).
 CC -1- ENZYME REGULATION: GLUCAN SYNTHESIS BY GTF-S IS INDEPENDENT OF
 CC PRIMER GLUCAN UNLIKE GTF-I.
 CC -1- DISEASE: DENTAL CARIES.
 CC -1- MISCELLANEOUS: GTF-S SYNTHESIZES WATER-SOLUBLE GLUCANS (ALPHA
 CC 1,6-GLUCOSE).
 CC -1- SIMILARITY: TO OTHER GLUCOSYLTRANSFERASES AND SOME TO A GLUCAN-
 CC BINDING PROTEIN FROM S. MUTANS.
 CC -1- SIMILARITY: Contains 10 cell wall binding repeats.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC EMBL; M30943; AAA26898.1; -

```

DR InterPro; IPR002479; CW binding.
DR InterPro; IPR003318; Glyco_hydro_70.
DR Pfam; PF01473; CW binding 1; 8.
DR Pfam; PF02324; Glyco_hydro_70; 1.
KW Transferrase; Glycosyltransferase; Signal; Repeat; Dental caries.
FT SIGNAL 1
FT CHAIN 37 1365 OR 37 (POTENTIAL).
FT REPEAT 157 177 CELL WALL BINDING 1.
FT REPEAT 178 197 CELL WALL BINDING 1.
FT REPEAT 198 197 CELL WALL BINDING 1.
FT DOMAIN 198 1061 CATALYTIC (APPROXIMATE).
FT REPEAT 1062 1082 CELL WALL BINDING 3.
FT REPEAT 1083 1102 CELL WALL BINDING 4.
FT REPEAT 1150 1169 CELL WALL BINDING 5.
FT REPEAT 1170 1190 CELL WALL BINDING 6.
FT REPEAT 1225 1243 CELL WALL BINDING 7.
FT REPEAT 1289 1308 CELL WALL BINDING 8.
FT REPEAT 1309 1328 CELL WALL BINDING 9.
FT REPEAT 1331 1352 CELL WALL BINDING 10.
FT REPEAT 1365 AA; 151590 MW; 167236B5A2EBC476 CRC64;
SQ SEQUENCE

Query Match 65.5%; Score 78; DB 1; Length 1365;
Best Local Similarity 71.4%; Pred. No. 0.00037;
Matches 15; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

OY 1 ANDHLSTLEKMSDNDTPYVLD 21
Db 467 AIDHLSTLEKMSGNDNDYVD 487

RESULT 6
GTFD_STRMU STANDARD; PRT; 1462 AA.
ID PF9331, O69383, O69386; O69389; O69392; O69398;
AC 01-FEB-1996 (Rel. 33, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DR 15-SEP-2003 (Rel. 42, Last annotation update)
DE Glucosyltransferase-S precursor (EC 2.4.1.5) (GTF-S) (Dextranucrase)
DE (sucrose 6-glucosyltransferase).
GN GTFD OR SMU.910.
OS Streptococcus mutans.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1309;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=GS-5;
RX MEDLINE=91100958; PubMed=2148600;
RA Honda O., Kato C., Kuramitsu H.K.;
RT "Nucleotide sequence of the Streptococcus mutans gtfD gene encoding
the glucosyltransferase-S enzyme."
RL J. Gen. Microbiol. 136:2099-2105(1990).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=MT4239 / Serotype C, MT4245 / Serotype E, MT4251 / Serotype F,
MT4467 / Serotype E, and MT8148 / Serotype C;
RX MEDLINE=98231643; PubMed=9570124;
RA Fujiwara T., Terzo Y., Hoshino T., Kawabata S., Ooshima T., Sobue S.,
Kimura S., Hamada S.;
RT "Molecular analyses of glucosyltransferase genes among strains of
Streptococcus mutans."
RL FEMS Microbiol. Lett. 161:331-336(1998).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=UAI59 / ATCC 700610 / Serotype C;
RX MEDLINE=22295063; PubMed=1397186;
RA Ajdacic D., McShan W.M., McLaughlin R.E., Savic G., Chang J., Qian Y.,
Carson M.B., Primeaux C., Tian R., Kenton S., Jia H., Lin S., Qian Y.,
Li S., Zhu H., Najjar F., Lai H., White J., Roe B.A., Ferretti J.J.;
RT "Genome sequence of Streptococcus mutans UAI59, a cariogenic dental
pathogen."
RL Proc. Natl. Acad. Sci. U.S.A. 99:14434-14439(2002).
CC -!- FUNCTION: PRODUCTION OF EXTRACELLULAR GLUCANS, THAT ARE THOUGHT
TO PLAY A KEY ROLE IN THE DEVELOPMENT OF THE DENTAL PLAQUE BECAUSE

```

```

CC OF THEIR ABILITY TO ADHERE TO SMOOTH SURFACES AND MEDIANE THE
CC AGGREGATION OF BACTERIAL CELLS AND FOOD DEBRIS.
CC -!- CATALYTIC ACTIVITY: Sucrose + ((1,6)-alpha-D-glucosyl)(N) = D-
CC fructose + ((1,6)-alpha-D-glucosyl)(N+1).
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- DISEASE: DENTAL CARIES.
CC -!- MISCELLANEOUS: GTF-1 SYNTHESIZES WATER-INSOLUBLE GLUCANS (ALPHA
CC 1,3-LINKED GLUCOSE AND SOME 1,6 LINKAGES), GTF-S SYNTHESIZES
CC WATER-SOLUBLE GLUCANS (ALPHA 1,6-GLUCOSE). GTF-SI SYNTHESIZES BOTH
CC FORMS OF GLUCANS.
CC -!- SIMILARITY: TO OTHER GLUCOSYLTRANSFERASES AND SOME TO A GLUCAN-
CC BINDING PROTEIN FROM S. MUTANS.
CC -!- SIMILARITY: Contains 6 cell wall binding repeats.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL; M29296; AAA26895.1; -
CC DR EMBL; D88653; BAA26103.1; -
CC DR EMBL; D88656; BAA26107.1; -
CC DR EMBL; D88659; BAA26111.1; -
CC DR EMBL; D88662; BAA26115.1; -
CC DR EMBL; D89979; BAA26121.1; -
CC DR EMBL; AE014932; AAN58619.1; -
CC DR InterPro; IPR002479; CW binding.
CC DR InterPro; IPR003318; Glyco_hydro_70.
CC DR Pfam; PF02324; Glyco_hydro_70; 1.
CC KW Transferrase; Glycosyltransferase; Signal; Repeat; Dental caries;
CC Complete proteome.
CC KM
CC FT CHAIN 1 1462 ?
CC FT SIGNAL 2 1462 1423
CC FT DOMAIN 1232 1295
CC FT REPEAT 1232 1295
CC FT REPEAT 1296 1359
CC FT REPEAT 1360 1423
CC FT REPEAT 10 10
CC FT VARIANT 19 19
CC FT VARIANT 58 58
CC FT VARIANT 68 68
CC FT VARIANT 81 81
CC FT VARIANT 113 113
CC FT VARIANT 122 122
CC FT VARIANT 132 132
CC FT VARIANT 135 135
CC FT VARIANT 137 137
CC FT VARIANT 202 202
CC FT VARIANT 255 255
CC FT VARIANT 275 275
CC FT VARIANT 288 288
CC FT VARIANT 301 301
CC FT VARIANT 313 313
CC FT VARIANT 317 317
CC FT VARIANT 328 328
CC FT VARIANT 350 350
CC FT VARIANT 628 633
CC FT VARIANT 658 668
CC FT VARIANT 726 732
CC FT VARIANT 726 730
CC FT VARIANT 762 762
CC FT VARIANT 762 762

```

```

POTENTIAL.
GLUCOSYLTRANSFERASE-S.
3 X 63 AA APPROXIMATE TANDEM REPEATS.
1.
2.
3.
Y -> H (IN STRAINS GS-5, MT4239, MT4245,
MT4251, MT4467 AND MT8148).
I -> V (IN STRAINS GS-5, MT4239, MT4245,
MT4251, MT4467 AND MT8148).
K -> E (IN STRAIN MT4467).
A -> S (IN STRAINS MT4239 AND MT4245).
A -> T (IN STRAINS MT4251 AND MT8148).
T -> I (IN STRAINS MT4239 AND MT4245).
A -> V (IN STRAINS MT4239, MT4245 AND
MT8148).
A -> S (IN STRAINS GS-5 AND MT4467).
A -> V (IN STRAIN MT4245).
A -> T (IN STRAINS GS-5, MT4239, MT4245,
MT4251, MT4467 AND MT8148).
V -> L (IN STRAIN MT4239).
D -> N (IN STRAIN MT8148).
E -> D (IN STRAINS MT4239, MT4245 AND
MT4251).
D -> N (IN STRAINS MT4239, MT4245 AND
MT4251).
Q -> H (IN STRAIN MT4245).
D -> N (IN STRAINS MT4239 AND MT4251).
E -> K (IN STRAIN MT4239).
V -> F (IN STRAIN MT4239).
F -> L (IN STRAINS MT4239, MT4251 AND
MT4467).
KKKYYTQ -> EKEYTL (IN STRAIN MT4251).
A -> S (IN STRAIN MT4239).
TDOGSEA -> ADKNGS (IN STRAIN MT4251).
TDQGS -> ADKGN (IN STRAINS MT4239 AND
MT4245).
T -> A (IN STRAINS GS-5, MT4239, MT4245,
MT4251, MT4467 AND MT8148).

```

```

FT  VARIANT  964  964  D -> Y (IN STRAIN MT4251)
FT  VARIANT  1019 1019  E -> K (IN STRAINS MT4245 AND MT4251)
FT  VARIANT  1059 1060  LG -> R (IN STRAIN MT4251)
FT  VARIANT  1060 1060  G -> R (IN STRAIN MT4245)
FT  VARIANT  1080 1080  H -> R (IN STRAIN MT4239)
FT  VARIANT  1142 1142  S -> Q (IN STRAIN GS-5)
FT  VARIANT  1198 1198  S -> N (IN STRAIN MT4251)
FT  VARIANT  1220 1220  Y -> C (IN STRAIN MT4251)
FT  VARIANT  1280 1280  F -> L (IN STRAIN MT4245)
FT  VARIANT  1282 1282  Q -> P (IN STRAIN MT4245)
FT  VARIANT  1290 1290  K -> T (IN STRAIN MT4245)
FT  VARIANT  1311 1311  N -> D (IN STRAIN GS-5)
FT  VARIANT  1403 1403  G -> R (IN STRAIN GS-5)
FT  VARIANT  1425 1425  R -> K (IN STRAIN MT4245)
FT  VARIANT  1449 1449  RYDKNSGNVYKRVTLANGRRITGIDRMGIARY -> VY
FT  CONFLICT 1428 1462  R (IN REF. 1)
SQ  SEQUENCE 1462 AA; 163387 MW; CEA4279CAD708645 CRC64;

Query Match 63.0%; Score 75; DB 1; Length 1462;
Best Local Similarity 71.4%; Pred. No. 0.0012;
Matches 15; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

Ox 1 ANDHSLTEAMSDNDPEYLD 21
Db 495 AINHLSTLEAMSDNDPQINKD 515

RESULT 7
G3PC_LEIME STANDARD; PRT; 330 AA.
ID G3PC_LEIME
AC Q01558;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Glyceroldehyde 3-phosphate dehydrogenase, cytosolic (EC 1.2.1.12)
DE (GAPDH)
DE GAPC.
OS Leishmania mexicana.
OC Eukaryota; Euklenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.
OX NCBI_Taxid=5665;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SSP. Mexicana;
RX MEDLINE=93063042; PubMed=1435864;
RA Hamaer V., Blaauw M., Kohl L., Allert S., Oppertoes F.R.,
RA Michels P.A.M.;
RT "Molecular analysis of the cytosolic and glycosomal glyceroldehyde-3-
RT phosphate dehydrogenase in Leishmania mexicana.";
RT Mol. Biochem. Parasitol. 55:115-126(1992).
CC -1- CATALYTIC ACTIVITY: D-glyceroldehyde 3-phosphate + phosphate +
CC NAD(+) = 3-phospho-D-glyceroyl phosphate + NADH.
CC -1- PATHWAY: Second phase of glycolysis; first step.
CC -1- SUBUNIT: Homotetramer.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: Belongs to the glyceroldehyde 3-phosphate
CC dehydrogenase family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: X65220; CAA46323.1; -
CC FTR; B48445; B48445.
CC HSP; P06977; IGAD.
CC InterPro: IPR000173; GAP_dhdtrogenase.
CC InterPro: IPR006424; GAPDH-I.
CC Pfam: PF00044; gpdh. 1.
CC Pfam: PF02800; gpdh_C. 1.

```

```

DR PRINTS; PR00078; G3PDHGRNASE.
DR TIGRfam; TIGR01534; GAPDH-1; 1.
DR PROSITE; PS00071; GAPDH; 1.
KW Glycolysis; Oxidoreductase; NAD.
FT INIT_MET 0 BY SIMILARITY.
FT BINDING 148 148 GLYCERALDEHYDE 3-PHOSPHATE
FT ACT_SITE 175 175 ACTIVATES THIOL GROUP DURING CATALYSIS.
SQ SEQUENCE 330 AA; 35511 MW; ED4B6DBE8A207F1E CRC64;

Query Match 47.9%; Score 57; DB 1; Length 330;
Best Local Similarity 52.6%; Pred. No. 0.14;
Matches 10; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

Ox 2 NDHSLTEAMSDNDPEYLD 20
Db 300 NDHFLVTLVSWYDNETGYSH 318

RESULT 8
YBHK_SALTI STANDARD; PRT; 302 AA.
ID YBHK_SALTI
AC P58587;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Hypothetical UPF0052 protein ybhK.
DE YBHK OR STY0835 OR T2086.
GN YBHK_SALTI.
OS Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Salmonella.
OX NCBI_Taxid=601;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CT18;
RX MEDLINE=21534947; PubMed=11677608;
RA Parkhill J., Dougan G., James K.D., Thomson N.R., Pickard D., Wain J.,
RA Churcher C., Mungall K.L., Bentley S.D., Holden M.T.G., Sebatina M.,
RA Baker S., Basham D., Brooks K., Chillingworth T., Connor P.,
RA Cronin A., Davis P., Davies R.M., Dowd L., White N., Farrar J.,
RA Fellwell T., Hamlin N., Haque A., Hien T.T., Holtroyd S., Jagers K.,
RA Krogan A., Larsen T.S., Leather S., Moule S., O'Gaora P., Parry C.,
RA Quail M., Rutherford K., Simmonds M., Skelton J., Stevens K.,
RA Whitehead S., Barrett B.G.;
RT "Complete genome sequence of a multiple drug resistant Salmonella
RT enterica serovar Typhi CT18.";
RT Nature 413:848-852(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Ty2 / ATCC 700931;
RX MEDLINE=22531367; PubMed=12644504;
RA Deng W., Lion S.-R., Plunkett G. III, Mayhew G.F., Rose D.J.,
RA Burland V., Kodyanski V., Schwartz D.C., Blatner F.R.,
RT "Comparative genomics of Salmonella enterica serovar Typhi strains Ty2
RT and CT18";
RT J. Bacteriol. 185:2330-2337(2003).
CC -1- SIMILARITY: Belongs to the UPF0052 family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AL627268; CAD05249.1; -
CC EMBL: AE015841; AAC069704.1; -
CC InterPro: IPR002882; UPF0052.
CC Pfam: PF01933; UPF0052. 1.
CC Hypothetical protein; Complete proteome.
SQ SEQUENCE 302 AA; 32656 MW; 221847121E185C3 CRC64;

Query Match 41.2%; Score 49; DB 1; Length 302;

```

Best Local Similarity 50.0%; Pred. No. 2.4;
Matches 10; Conservative 1; Mismatches 9; Indels 0; Gaps 0;
QY 2 NDHLSILEAMSDNDTPYHLD 21
DB 267 NDRLVIOEVLASDIPYRHD 286

RESULT 9
YBHK_SALTY STANDARD; PRT; 302 AA.
AC P58586; 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Hypothetical UPF0052 protein ybhK.
GN YBHK OR STM0801.
OS Salmonella typhimurium.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Salmonella.
OX NCBI_TaxId=602;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=UT2 / SCS01412 / ATCC 700720;
RX MEDLINE=21534948; PubMed=1677609;
RA McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,
RA Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D.,
RA Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,
RA Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,
RA Waterston R., Wilson R.K.; Miller W., Stoneking T., Nhan M.,
RT "Complete genome sequence of Salmonella enterica serovar Typhimurium
LT2.";
RL Nature 413:852-856(2001).
CC -1- SIMILARITY: Belongs to the UPF0052 family.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: AE008733; AAL19738.1; -
DR StryGene; SG7272; ybhK.
DR InterPro: IPR002882; UPF0052.
DR Pfam: PF01933; UPF0052.1
KM Hypothetical protein; Complete proteome.
SQ SEQUENCE 302 AA; 32626 MW; 38A84708A1E185C3 CRC64;
Query Match 41.2%; Score 49; DB 1; Length 302;
Best Local Similarity 50.0%; Pred. No. 2.4;
Matches 10; Conservative 1; Mismatches 9; Indels 0; Gaps 0;
QY 2 NDHLSILEAMSDNDTPYHLD 21
DB 267 NDRLVIOEVLASDIPYRHD 286

RP SEQUENCE FROM N.A.
RX MEDLINE=84270667; PubMed=6087149;
RA Baer R., Bankier A.T., Biggin M.D., Deininger P.L., Farrell P.U.,
RA Gibson T.J., Hatfull G., Hudson G.S., Satchwell S.C., Seguin C.,
RA Tufnell P.S., Watfull G., Wilson R.K.; Miller W., Stoneking T., Nhan M.,
RT "DNA sequence and expression of the B95-8 Epstein-Barr virus genome.";
RL Nature 310:207-211(1984).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=92113548; PubMed=1662696;
RA Chen M.R., Hsu T.Y., Lin S.W., Chen J.Y., Yang C.S.;
RT "Cloning and characterization of cDNA clones corresponding to
RT transcripts from the BamHI G region of the Epstein-Barr virus genome
RT and expression of BGLF2.";
RL J. Gen. Virol. 72:3047-3055(1991).
CC -1- SIMILARITY: BELONGS TO FAMILY THAT GROUPS TOGETHER HSV-1 UL16,
CC HSV-6 ORF11R, EBV-1 46, HCMV UL94, EBV BDLF2, AND VZV 44.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: M60514; AAA5871.1; -
DR EMBL: V01555; CA24831.1; -
DR EMBL: S77132; AAB2113.1; -
DR PIR: C43044; Q0BE40.
DR InterPro: IPR004286; UL16_UL94.
DR Pfam: PF03044; UL16_UL94.1.
KM Late protein.
SQ SEQUENCE 336 AA; 36888 MW; 840937A15D5584C CRC64;
Query Match 41.2%; Score 49; DB 1; Length 336;
Best Local Similarity 41.2%; Pred. No. 2.8;
Matches 7; Conservative 6; Mismatches 4; Indels 0; Gaps 0;
QY 1 ANDHLSILEAMSDNDTP 17
DB 216 AGAHVNLIRGWTEDDSDP 232

RESULT 11
G3PC_TRYBB STANDARD; PRT; 330 AA.
AC P10097;
ID G3PC_TRYBB
DT 01-MAY-1989 (Rel. 10, Created)
DT 01-MAY-1992 (Rel. 22, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Glyceraldehyde 3-phosphate dehydrogenase, cytosolic (BC 1.2.1.12)
DE (GAPDH).
OS Trypanosoma brucei brucei.
OC Eukaryota; Eucelozoa; Kinetoplastida; Trypanosomatidae; Trypanosoma.
OX NCBI_TaxId=5702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=427;
RX MEDLINE=91249838; PubMed=2040303;
RA Michels P.A.M., Marchand M., Kohl L., Allert S., Wierenga R.K.,
RA Oppenoes F.R.;
RT "The cytosolic and glycosomal isoenzymes of glyceraldehyde-3-phosphate
RT dehydrogenase in Trypanosoma brucei have a distant evolutionary
RT relationship.";
RL Eur. J. Biochem. 198:421-428(1991).
RN [2]
RP PRELIMINARY SEQUENCE OF 1-85.
RX MEDLINE=87161817; PubMed=3830153;
RA Misset O., van Beuningen J., Lambel A.M., van der Meer R.,
RA Oppenoes F.R.;
RT "Glyceraldehyde-phosphate dehydrogenase from Trypanosoma brucei.
RT Comparison of the glycosomal and cytosolic isoenzymes.";
RN [1]

```

RL Eur. J. Biochem. 162:501-507 (1987).
CC -1- CATALYTIC ACTIVITY: D-glyceraldehyde 3-phosphate + phosphate +
CC NAD(+) = 3-phospho-D-glyceroyl phosphate + NADH.
CC -1- PATHWAY: Second phase of glycolysis; first step.
CC -1- SUBUNIT: Homotetramer.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: Belongs to the glyceraldehyde 3-phosphate
CC dehydrogenase family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X53472; CAA37568.1; -.
DR PIR: S16091; DEUTGC.
DR HSP: P06977; 1GAD.
DR InterPro: IPR00173; GAP_dhydrogenase.
DR InterPro: IPR006424; GAPDH-I.
DR Pfam: PF00044; gpdh. 1.
DR Pfam: PF02800; gpdh.C; 1.
DR PRINTS: PR00078; G3PDHRCNASE.
DR TIGRFS: TIGR01534; GAPDH-I; 1.
DR PROSITE: PS00071; GAPDH; 1.
DR KEGG: glycolysis; Oxidoreductase; NAD.
DR INIT MET 0
FT BINDING 149 149 GLYCERALDEHYDE 3-PHOSPHATE.
FT ACT SITE 176 176 ACTIVATES THIOL GROUP DURING CATALYSIS.
FT VARIANT 72 72 R -> I.
FT CONFLICT 56 56 A -> T (IN REF. 2).
FT CONFLICT 60 60 Q -> K (IN REF. 2).
SQ SEQUENCE 330 AA; 35503 MW; 7E1F0AFD4091B7C5 CRC64;

Query Match 39.9%; Score 47.5; DB 1; Length 330;
Best Local Similarity 43.5%; Pred. No. 4.7;
Matches 10; Conservative 4; Mismatches 6; Indels 3; Gaps 1;

QY 2 NDHLSILEAMSDNDTPY---LHD 21
Db 301 NDHFWKLVSWYDNETGYSNKVDH 323

RESULT 12
G3P COCPO STANDARD; PRT; 337 AA.
AC 06JTH3;
DT 15-SEP-2003 (Rel. 42, Created)
DT 15-SEP-2003 (Rel. 42, Last sequence update)
DE Glyceraldehyde 3-phosphate dehydrogenase (EC 1.2.1.12) (GAPDH).
GN Coccidioides posadasii.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Omygenales; mitosporic Omygenales; Coccidioides.
OX NCBI_TaxID=19306;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=C735;
RA Hung C.-Y., Yu J.-J., Cole G.-T.;
RT "Isolation and evaluation of expression of the glyceraldehyde-3-
RT phosphate dehydrogenase gene of Coccidioides immitis.";
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases
CC -1- CATALYTIC ACTIVITY: D-glyceraldehyde 3-phosphate + phosphate +
CC NAD(+) = 3-phospho-D-glyceroyl phosphate + NADH.
CC -1- PATHWAY: Second phase of glycolysis; first step.
CC -1- SUBUNIT: Homotetramer (By similarity).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: Belongs to the glyceraldehyde 3-phosphate
CC dehydrogenase family.
CC -----

```

```

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AF288134; AAN76496.1; -.
DR InterPro: IPR006424; GAPDH-I.
DR InterPro: IPR000173; GAP_dhydrogenase.
DR Pfam: PF00044; gpdh. 1.
DR Pfam: PF02800; gpdh.C; 1.
DR PRINTS: PR00078; G3PDHRCNASE.
DR TIGRFS: TIGR01534; GAPDH-I; 1.
DR PROSITE: PS00071; GAPDH; 1.
DR KEGG: glycolysis; Oxidoreductase; NAD.
FT BINDING 151 151 GLYCERALDEHYDE 3-PHOSPHATE.
FT ACT SITE 178 178 ACTIVATES THIOL GROUP DURING CATALYSIS.
SQ SEQUENCE 337 AA; 36400 MW; 62EA834CA536DAF2 CRC64;

Query Match 39.5%; Score 47; DB 1; Length 337;
Best Local Similarity 52.9%; Pred. No. 5.7;
Matches 9; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 2 NDHLSILEAMSDNDTPY 18
Db 303 NDHFWKLVSWYDNETGYSNKVDH 319

RESULT 13
SC35 YEAST STANDARD; PRT; 275 AA.
ID SC35 YEAST
AC P53271;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE SEC35 protein.
GN SEC35 OR YGR120C OR G6324 (Baker's yeast).
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=S288c / FY1679;
RA MEDLINE=97197982; PubMed=9046098;
RA van Dyck L., Tettelein H., Purnelle B., Goffeau A.;
RT "An 18.3 kb DNA fragment from yeast chromosome VII carries four
RT unknown open reading frames, the gene for an Aen synthase, remnants
RT of ty and three tRNA genes.";
RT Yeast 13:171-176(1997).
RN [2]
RP FUNCTION, AND SUBCELLULAR LOCATION.
RA MEDLINE=98270888; PubMed=9606204;
RA Vanheenen S.M., Cao X., Lupashin V.V., Barlowe C., Waters M.G.;
RT "Sec35p, a novel peripheral membrane protein, is required for ER to
RT Golgi vesicle docking.";
RT J. Cell Biol. 141:1107-1119(1998).
RN [3]
RP SUBUNIT.
RA MEDLINE=21563418; PubMed=11703943;
RA Whyte J.R., Munro S.;
RT "The Sec34/35 Golgi transport complex is related to the exocyst,
RT defining a family of complexes involved in multiple steps of membrane
RT traffic.";
RL Dev. Cell 1:527-537(2001).
CC -1- FUNCTION: Required for ER to Golgi vesicle docking. Not essential
CC for viability.
CC -1- SUBUNIT: Component of the Sec34/Sec35 complex which consists of
CC eight different proteins.
CC -1- SUBCELLULAR LOCATION: Membrane-associated.
CC -----

```


This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; X83099; CAA58155.1; -
EMBL; Z72905; CAA97130.1; -
PIR; S55978; S55978.
SGD; S0003352; SEC35.
GO; GO:0017119; C:Golgi transport complex; IPI.
GO; GO:0000300; C:peripheral membrane protein of membrane fra. .; IDA.
GO; GO:0005625; C:soluble fraction; IDA.
GO; GO:0005515; F:protein binding activity; IPI.
GO; GO:0006888; P:ER to Golgi transport; IGI.
GO; GO:0000301; P:retrograde (vesicle recycling within Golgi) . .; IMP.
Transport; Protein transport; Golgi stack; Membrane.
SEQUENCE 275 AA; 31799 MW; AA102D086FF3PAD7 CRC64;

Query Match 38.7%; Score 46; DB 1; Length 275;
Best Local Similarity 50.0%; Pred. No. 6.5;
Matches 7; Conservative 6; Mismatches 1; Indels 0; Gaps 0;

7 ILEAMSNDTPTLH 20
|||:|||||:
217 ILESCADSNSPYIH 230

RESULT 14
G3P CHLPN STANDARD; PRT; 335 AA.
ID G3P CHLPN
AC Q9Z7T0; Q9ZQHT;
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Glyceraldehyde 3-phosphate dehydrogenase (EC 1.2.1.12) (GAPDH).
GN GAP OR GAPB OR CPN0624 OR CP0123
OS Chlamydia pneumoniae (Chlamydia pneumoniae).
OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia.
NCBI_TaxID=83558;
RN NCBI_TaxID=83558;
RP SEQUENCE FROM N.A.
RC STRAIN=CHW029;
RX MEDLINE=9920606; PubMed=10192388;
RA Kaiman S., Mitchell W., Marathe R., Lamme C., Fan J., Hyman R.W.,
RA Olinger L., Grimwood J., Davis R.W., Stephens R.S.;
RT "Comparative genomes of Chlamydia pneumoniae and C. trachomatis.";
RL Nat. Genet. 21:385-389(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=AR39;
RX MEDLINE=20150255; PubMed=10684935;
RA Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.F.,
RA White O., Hickey E.K., Peterson J., Uitterback T., BERRY K., Bess S.,
RA Linher K., Weidman J., Khouri H., Craven B., Bowman C., Dodson R.,
RA Gwynn M., Nelson W., DeBoy R., Kolonay J., McClarty G., Salzberg S.L.,
RA Eisen J., Fraser C.M.;
RT "Genome sequences of Chlamydia trachomatis MoPn and Chlamydia pneumoniae AR39.";
RL Nucleic Acids Res. 28:1397-1406(2000).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=J138;
RX MEDLINE=20330349; PubMed=10871362; Tabuchi M., Kishi F., Ouchi K.,
RA Shirai M., Hirakawa H., Kimoto M., Tabuchi M., Kishi F., Ouchi K.,
RA Shiba T., Ishii K., Hattori M., Kuhara S., Nakazawa T.;
RT "Comparison of whole genome sequences of Chlamydia pneumoniae J138 from Japan and CHW029 from USA.";
RL Nucleic Acids Res. 28:2311-2314(2000).
CC -1- CATALYTIC ACTIVITY: D-glyceraldehyde 3-phosphate + phosphate + NAD(+)= 3-phospho-D-glyceroyl phosphate + NADH.

-1- PATHWAY: Second phase of glycolysis; first step.
-1- SUBUNIT: Homotrimer (By similarity).
-1- SUBCELLULAR LOCATION: Cytoplasmic.
-1- SIMILARITY: Belongs to the glyceraldehyde 3-phosphate dehydrogenase family.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; AE01647; AAD18763.1; -
EMBL; AE002173; AAB38006.1; -
EMBL; AF002547; BAA98831.1; -
PIR; B72053; B72053.
PIR; B86568; B86568.
HSSP; P06977; IGAD.
PDB; P06977; 1GAD.
PDB; P06977; 1GAD.
PDB; P06977; 1GAD.
TIGR; CP0123; -
InterPro: IPR000173; GAP dehydrogenase.
InterPro: IPR006424; GAPDH-1.
Pfam: PF00044; spdh.1.
Pfam: PF02800; spdh.C.1.
PRINTS; PR00078; G3PDHGNASE.
TIGRFAMS; TIGR01534; GAPDH-1; 1.
PROSITE; PS00071; GAPDH; 1.
Glycolysis; Oxidoreductase; NAD; Complete proteome.
GLYCERALDEHYDE 3-PHOSPHATE (BY SIMILARITY).
ACTIVATES THIOL GROUP DURING CATALYSIS (BY SIMILARITY).
ACT_SITE 178 178
SEQUENCE 335 AA; 36837 MW; C86D0E3AD3ADDF8E CRC64;

Query Match 38.7%; Score 46; DB 1; Length 335;
Best Local Similarity 52.9%; Pred. No. 8.2;
Matches 9; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

2 NDHLSILEAMSNDTPTY 18
|||:|||||:
303 NDHFPFLVAMWYDNEYGY 319

RESULT 15
PCHA_PSEAE STANDARD; PRT; 476 AA.
ID PCHA_PSEAE
AC Q51506;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Salicylate biosynthesis isochorismate synthase (EC 5.4.99.6) (isochorismate mutase).
GN PCHA OR PA4231.
OS Pseudomonas aeruginosa.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales; Pseudomonadaceae; Pseudomonas.
NCBI_TaxID=287;
RN NCBI_TaxID=287;
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 15692 / PAO1;
RX MEDLINE=9608693; PubMed=7500944;
RA Serrino L., Reilmann C., Baur H., Beyeler M., Visca P., Haas D.;
RT "Structural genes for salicylate biosynthesis from chorismate in Pseudomonas aeruginosa.";
RL Mol. Gen. Genet. 249:217-228(1995).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 15692 / PAO1;
RX MEDLINE=20437337; PubMed=10984043;
RA Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warrenner P., Hickey M.J., Brinkman F.S.L., Hutnagle W.O., Kowalik D.J., Lagrou M.,


```

RA Garber R.L., Goltzy L., Tolentino E., Westbrock-Wadman S., Yuan Y.,
RA Brody L.J., Coulter S.N., Folger K.R., Kas A., Lardig K., Lim R.M.,
RA Smith K.A., Spencer D.H., Mong G.K.-S., Wu Z., Paulsen I.T.,
RA Reizer J., Sailer M.H., Hancock R.E.W., Lory S., Olson M.V.,
RA "Complete genome sequence of Pseudomonas aeruginosa PA01, an
RT opportunistic pathogen." ;
RL Nature 406:959-964(2000).
CC -1- FUNCTION: INVOLVED IN THE CONVERSION OF CHORISMATE TO SALICYLATE
CC -1- (PROBABLE).
CC -1- CATALYTIC ACTIVITY: Chorismate = isochorismate.
CC -1- PATHWAY: Salicylic acid biosynthesis.
CC -1- SIMILARITY: BELONGS TO THE ISOCHORISMATE SYNTHASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as their content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@ib-sib.ch).
CC -----
DR EMBL: X82644; CNA57969.1; -.
DR EMBL: AE004840; AAG07619.1; -.
DR PIR: S6203; S58229.
DR InterPro: IPR005801; Anth_synth_chor.
DR InterPro: IPR004561; Isochor_syn.
DR Pfam: PF00425; chorismate_bind.1.
DR ProDom: PD000779; Anth_synth_chor; 1.
DR TIGRFAMs: TIGR00543; isochor_syn; 1.
KW Isomerase; Complete proteome.
SQ SEQUENCE 476 Aa; 52071 Mw; D0F6AFP9DCDF9CF5 CRC64;

Query Match 38.7%; Score 46; DB 1; Length 476;
Best Local Similarity 44.4%; Pred. No. 13;
Matches 8; Conservative 2; Mismatches 8; Indels 0; Gaps 0.

3 DDLSTLEANSNDNDPTLYH 20
44 DPLQVFGANDDKOTPCPLY 61

```

Search completed: November 13, 2003, 09:45:25
Job time : 7.56398 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 13, 2003, 09:31:40 ; Search time 31.6493 Seconds
(without alignments)
171.224 Million cell updates/sec

Title: US-09-290-049A-10
Perfect score: 119
Sequence: 1 ANDHLSIEFMSDNDPTPLHD 21

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

SPTREMBL.23:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phase:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteriaph:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	108	90.8	1590	2	Q59983 streptococc
2	108	90.8	1590	2	Q55263 streptococc
3	78	65.5	1338	2	Q9WXJ4 streptococc
4	76	63.9	1016	2	Q9LCT7 leucostoc
5	74	62.2	1477	2	Q91466 leucostoc
6	74	62.2	1508	2	Q9EZHS leucostoc
7	74	62.2	1508	2	Q52224 leucostoc
8	73	61.3	1575	2	Q9LCH3 streptococc
9	73	61.3	1577	2	Q54178 streptococc
10	71	59.7	1527	2	Q9ZAR4 leucostoc
11	71	59.7	1527	2	Q8KRE1 leucostoc
12	69	58.0	2835	2	Q8G9Q2 streptococc
13	66	55.5	1449	2	Q68542 streptococc
14	66	55.5	1449	2	Q55264 streptococc
15	63	52.9	1554	2	Q8KZL5 streptococc
16	63	52.9	1577	2	Q55265 streptococc

17	62	52.1	1512	2	Q9WXJ5	Q9WXJ5 streptococc
18	60	50.4	2057	2	Q9RE05	Q9RE05 leucostoc
19	58	48.7	1518	2	Q00600	Q00600 streptococc
20	57	47.9	1599	2	Q00599	Q00599 streptococc
21	55.5	46.6	282	10	Q8RWY3	Q8RWY3 allium cepa
22	52	43.7	4848	2	Q07944	Q07944 streptococc
23	51	42.9	424	10	Q94U07	Q94U07 arabidopsis
24	51	42.9	435	10	Q9L068	Q9L068 arabidopsis
25	50	42.0	305	10	Q8LDL1	Q8LDL1 arabidopsis
26	50	42.0	339	10	Q949T9	Q949T9 arabidopsis
27	50	42.0	347	10	Q9LY92	Q9LY92 arabidopsis
28	49.5	41.6	55	10	Q94FX3	Q94FX3 zea mays
29	49.5	41.6	55	10	Q93VN4	Q93VN4 zea mays
30	49.5	41.6	55	10	Q93WH7	Q93WH7 zea mays
31	49.5	41.6	56	10	Q94FX8	Q94FX8 zea mays
32	49.5	41.6	129	10	Q94EX6	Q94EX6 zea mays
33	49.5	41.6	547	10	Q8LST4	Q8LST4 sorghum bic
34	49.5	41.6	549	10	Q94G64	Q94G64 zea mays
35	49.5	41.6	549	10	Q43274	Q43274 zea mays
36	49.5	41.6	601	10	Q05566	Q05566 mycobacteri
37	49	41.2	797	16	Q98P32	Q98P32 rhizobium l
38	48.5	40.8	283	5	Q9VLP7	Q9VLP7 dirosophila
39	48.5	40.8	283	5	Q8WSF4	Q8WSF4 dirosophila
40	48	40.3	490	16	Q06579	Q06579 mycobacteri
41	47.5	39.9	214	11	Q8C197	Q8C197 mus musculu
42	47.5	39.9	519	10	Q9ZUB6	Q9ZUB6 arabidopsis
43	47.5	39.9	534	10	Q8S528	Q8S528 arabidopsis
44	47.5	39.9	534	10	Q94C67	Q94C67 arabidopsis
45	47.5	39.9	549	10	Q91LR2	Q91LR2 oryza sativ

ALIGNMENTS

RESULT 1

ID Q59983 PRELIMINARY; PRT; 1590 AA.
AC Q59983;
DC 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-OCT-2002 (TREMBLrel. 22, Last annotation update)
DE Glucosyltransferase-I precursor (EC 2.4.1.5).
GN Grp1.
OS Streptococcus sobrinus.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
CC Streptococcus.
CX NCBI_TaxID=1310;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=OM2176;
RX MEDLINE=9416405; PubMed=8312602;
RA Sato S., Inoue M., Handa N., Aizawa Y., Isobe Y., Katayama T.;
RT "DNA sequence of the glucosyltransferase gene of serotype d
RT Streptococcus sobrinus.";
RL DNA Seq. 4:19-27(1993).
DR EMBL; D13858; BAA02976.1; .
DR InterPro; IPR003118; Glyco_hydro_70.
DR Pfam; PF01473; CM binding_1; 16.
DR Pfam; PF03324; Glyco_hydro_70; 1.
KW Glycosyltransferase; Signal; Transferase.
FT SIGNAL
FT CHAIN 1
FT SIGNAL 38
SQ SEQUENCE 1590 AA; 175955 MW; C3C83A57CF3C2B0E CRC64;

Query Match 90.8%; Score 108; DB 2; Length 1590;
Best Local Similarity 85.7%; Pred. No. 5.5e-08;
Matches 18; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Cy 1 ANDHLSIEFMSDNDPTPLHD 21
Db 477 ANNHSLIEFMSDNDPTPLHD 497

RESULT 2

Q55263 PRELIMINARY; PRT; 1590 AA.
 AC Q55263; (TREMBlrel. 01, Created)
 DT 01-NOV-1996 (TREMBlrel. 05, Last sequence update)
 DT 01-NOV-1998 (TREMBlrel. 22, Last annotation update)
 DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
 DE GTF-1.
 GN GLUCOSYLTRANSFERASE.
 OS Streptococcus sobrinus.
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 NCBI_TaxID=1310;
 RN NCBI_TaxID=1310;
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 33478;
 RA Sato S.;
 RT "DNA and amino-acid sequences of water-insoluble-glucan synthetase produced from Streptococcus sobrinus ATCC 33478.";
 RL Am. Kagoshima Univ. Dental School 16:23-29(1996).
 DR EMBL; D63570; BAA09792.1; -;
 DR InterPro; IPR002479; CW_binding.
 DR InterPro; IPR003318; Glyco_hydro_70.
 DR Pfam; PF01473; CW_binding_1; 15.
 DR Pfam; PF02324; Glyco_hydro_70; 1.
 SQ SEQUENCE 1590 AA; 176058 MW; 9DF7A3F2C6B4FD43 CRC64;

Query Match 90.8%; Score 108; DB 2; Length 1590;
 Best Local Similarity 85.7%; Pred. No. 5.5e-08;
 Matches 18; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 ANDHSLILEAWSNDNTPYLHD 21
 DB 477 ANNHSLILEAWSNDNTPYLHD 497

RESULT 3

Q9WXJ4 PRELIMINARY; PRT; 1338 AA.
 AC Q9WXJ4; (TREMBlrel. 12, Created)
 DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
 DT 01-NOV-1999 (TREMBlrel. 22, Last annotation update)
 DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
 DE GTF-8.
 GN Streptococcus criceti.
 OS Streptococcus criceti.
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 NCBI_TaxID=1333;
 RN NCBI_TaxID=1333;
 RP SEQUENCE FROM N.A.
 RC STRAIN=HS-6;
 RA Inoue M., Fukui K., Miyagi A.;
 RT "Sequence analysis of the gene encoding a glucosyltransferase (gtrf) genes.";
 RL Submitted (MAR-1999) to the EMBL/Genbank/DBJ databases.
 DR EMBL; AB026123; BAA77236.1; -;
 DR InterPro; IPR002479; CW_binding.
 DR InterPro; IPR003318; Glyco_hydro_70.
 DR Pfam; PF01473; CW_binding_1; 10.
 DR Pfam; PF02324; Glyco_hydro_70; 1.
 KW plasmid.
 SQ SEQUENCE 1338 AA; 148558 MW; 0A90C8E10E15D99B CRC64;

Query Match 65.5%; Score 78; DB 2; Length 1338;
 Best Local Similarity 71.4%; Pred. No. 0.0025;
 Matches 15; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 ANDHSLILEAWSNDNTPYLHD 21
 DB 437 AIDHSLILEAWSNDNTPYLHD 457

RESULT 4

Q9LCJ7 PRELIMINARY; PRT; 1016 AA.
 AC Q9LCJ7; (TREMBlrel. 15, Created)
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
 DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
 DE Dextranase.
 GN DSR.
 OS Leuconostoc mesenteroides.
 OC Bacteria; Firmicutes; Lactobacillales; Leuconostoc.
 NCBI_TaxID=1245;
 RN NCBI_TaxID=1245;
 RP SEQUENCE FROM N.A.
 RC STRAIN=NRRL B-512F;
 RX MEDLINE=20169623; PubMed=10705445;
 RA Funari K., Mizuno K., Takahara H., Kobayashi M.;
 RT "Gene encoding a dextranase-like protein in Leuconostoc mesenteroides NRRL B-512F.";
 RL Biochem. Biotechnol. 64:29-38(2000).
 DR EMBL; AB020020; BAA90527.1; -;
 DR HSP; P06278; 1VUS.
 DR InterPro; IPR003318; Glyco_hydro_70.
 DR Pfam; PF02324; Glyco_hydro_70; 1.
 SQ SEQUENCE 1016 AA; 110344 MW; 8896EFDE13CCCB47 CRC64;

Query Match 63.9%; Score 76; DB 2; Length 1016;
 Best Local Similarity 61.9%; Pred. No. 0.0037;
 Matches 13; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 ANDHSLILEAWSNDNTPYLHD 21
 DB 548 SNOHSLILEAWSNDNTPYLHD 568

RESULT 5

Q9L466 PRELIMINARY; PRT; 1477 AA.
 AC Q9L466; (TREMBlrel. 15, Created)
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
 DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
 DE Dextranase (EC 2.4.1.5).
 GN DSR.
 OS Leuconostoc mesenteroides.
 OC Bacteria; Firmicutes; Lactobacillales; Leuconostoc.
 NCBI_TaxID=1245;
 RN NCBI_TaxID=1245;
 RP SEQUENCE FROM N.A.
 RC STRAIN=NRRL B-1355;
 RA Arguello-Morales M.A., Renaud-Simeon M., Pizzut S., Sarcabal P.,
 RA Willenot R.M., Monsen P.;
 RT "Sequence analysis of the gene encoding a dextranase, a sucrose
 RT glucosyltransferase from Leuconostoc mesenteroides NRRL B-1355.";
 RL Submitted (OCT-1999) to the EMBL/Genbank/DBJ databases.
 DR EMBL; AJ250172; CAB76565.1; -;
 DR InterPro; IPR002479; CW_binding.
 DR InterPro; IPR003318; Glyco_hydro_70.
 DR Pfam; PF01473; CW_binding_1; 14.
 DR Pfam; PF02324; Glyco_hydro_70; 1.
 KW Glycosyltransferase; Transferase.
 SQ SEQUENCE 1477 AA; 164887 MW; E6F5710DEDFC831 CRC64;

Query Match 62.2%; Score 74; DB 2; Length 1477;
 Best Local Similarity 66.7%; Pred. No. 0.012;
 Matches 14; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 1 ANDHSLILEAWSNDNTPYLHD 21
 DB 532 ANGHSLILEAWSNDNTPYLHD 552

RESULT 6
O9EZHS PRELIMINARY; PRT: 1508 AA.
ID O9EZHS
AC O9EZHS
DT 01-MAR-2001 (TRENBLrel. 16, Created)
DT 01-MAR-2001 (TRENBLrel. 16, Last sequence update)
DT 01-OCT-2002 (TRENBLrel. 22, Last annotation update)
DE Dextranucrase Dszb742.
GN DSRB742.
OS Leuconostoc mesenteroides.
OC Bacteria; Firmicutes; Lactobacillales; Leuconostoc.
OX NCBI_TaxID=1245;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=B-742CB;
RA Kim H.-S., Kim D., Ryu H.-J., Robyt J.F.;
RT "Leuconostoc mesenteroides B-742CB, a dextranucrase gene."
RL Submitted (Aug-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF294469; AAG38021.1; -
DR InterPro; IPR002479; CW_binding.
DR InterPro; IPR003318; Glyco_hydro_70.
DR Pfam; PF01473; CW_binding_1; 14.
DR Pfam; PF02324; Glyco_hydro_70; 1.
SQ SEQUENCE 1508 AA; 168542 MW; 1.E2FCFA0F87AE4F3A CRC64;
KW Transferase.
Query Match 62.2%; Score 74; DB 2; Length 1508;
Best Local Similarity 66.7%; Pred. No. 0.012;
Matches 14; Conservative 1; Mismatches 6; Indels 0; Gaps 0;
QY 1 ANDHLSLEAWSNDPTPLHD 21
Db 563 ANQHLSTLEDSHNDPEYVKD 583
RESULT 7
O52224 PRELIMINARY; PRT: 1508 AA.
ID O52224
AC O52224
DT 01-JUN-1998 (TRENBLrel. 06, Created)
DT 01-JUN-1998 (TRENBLrel. 06, Last sequence update)
DT 01-OCT-2002 (TRENBLrel. 22, Last annotation update)
DE Glucosyltransferase (EC 2.4.1.5).
GN DSRB.
OS Leuconostoc mesenteroides.
OC Bacteria; Firmicutes; Lactobacillales; Leuconostoc.
OX NCBI_TaxID=1245;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NRRL B-1299;
RA Monchois V., Renaud-Simeon M., Monsan P., Willemot R.M.;
RT "Cloning and sequencing of a gene coding for an extracellular
RT dextranucrase (DSRB) from Leuconostoc mesenteroides NRRL B-1299
RT synthesizing only a a(1-6) glucan."
RL FEMS Microbiol. Lett. 0:0-0(1998).
DR EMBL; AF030129; AAB95453.1; -
DR InterPro; IPR002479; CW_binding.
DR InterPro; IPR003318; Glyco_hydro_70.
DR Pfam; PF01473; CW_binding_1; 14.
DR Pfam; PF02324; Glyco_hydro_70; 1.
DR Glycosyltransferase; transferase.
SQ SEQUENCE 1508 AA; 168511 MW; E70CECB57A70D1F0 CRC64;
KW
Query Match 62.2%; Score 74; DB 2; Length 1508;
Best Local Similarity 66.7%; Pred. No. 0.012;
Matches 14; Conservative 1; Mismatches 6; Indels 0; Gaps 0;
QY 1 ANDHLSLEAWSNDPTPLHD 21
Db 563 ANQHLSTLEDSHNDPEYVKD 583

ID O9LCH3 PRELIMINARY; PRT: 1575 AA.
AC O9LCH3
DT 01-OCT-2000 (TRENBLrel. 15, Created)
DT 01-OCT-2000 (TRENBLrel. 15, Last sequence update)
DT 01-OCT-2002 (TRENBLrel. 22, Last annotation update)
DE Glucosyltransferase.
GN GTRF.
OS Streptococcus oralis.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OX Streptococcus.
OX NCBI_TaxID=1303;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC10557;
RX MEDLINE=20231779; PubMed=10768934;
RA Fujiwara T., Hoshino T., Ooshima T., Sobue S., Hamada S.;
RT "Purification, characterization, and molecular analysis of the gene
RT encoding glucosyltransferase from Streptococcus oralis."
RL Infect. Immun. 68:2475-2483(2000).
DR EMBL; AB025228; BAA95201.1; -
DR InterPro; IPR002479; CW_binding.
DR InterPro; IPR003318; Glyco_hydro_70.
DR Pfam; PF01473; CW_binding_1; 17.
DR Pfam; PF02324; Glyco_hydro_70; 1.
KW Transferase.
SQ SEQUENCE 1575 AA; 176792 MW; 772A26E4D7C2B543 CRC64;
KW
Query Match 61.3%; Score 73; DB 2; Length 1575;
Best Local Similarity 77.8%; Pred. No. 0.018;
Matches 14; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 4 HLSLEAWSNDPTPLHD 21
Db 549 HLSLEAWSNDPTPLHD 566
RESULT 9
O54178 PRELIMINARY; PRT: 1577 AA.
ID O54178
AC O54178
DT 01-NOV-1996 (TRENBLrel. 01, Created)
DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)
DT 01-OCT-2002 (TRENBLrel. 22, Last annotation update)
DE Glucosyltransferase.
GN GTRF.
OS Streptococcus gordonii Challis.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OX Streptococcus.
OX NCBI_TaxID=29390;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CHALLIS;
RX MEDLINE=96157084; PubMed=8586195;
RA Vickerman M.M., Sulavik M.C., Clewell D.B.;
RT "Molecular analysis of Streptococcus gordonii glucosyltransferase
RT phase variants."
RL Dev. Biol. Stand. 85:309-314(1995).
RN [2]
RP SEQUENCE OF 1-96 FROM N.A.
RC STRAIN=CHALLIS;
RX MEDLINE=92276337; PubMed=1534326;
RA Sulavik M.C., Tardif G., Clewell D.B.;
RT "Identification of a gene, rgs, which regulates expression of
RT glucosyltransferase and influences the spp phenotype of Streptococcus
RT gordonii Challis."
RL J. Bacteriol. 174:3577-3586(1992).
DR EMBL; U12643; AAG3483.1; -
DR EMBL; M89776; AAA26969.1; -
DR InterPro; IPR002479; CW_binding.
DR InterPro; IPR003318; Glyco_hydro_70.
DR Pfam; PF01473; CW_binding_1; 18.
DR Pfam; PF02324; Glyco_hydro_70; 1.
KW Transferase.

SQ SEQUENCE 1577 AA; 177805 MW; 5AE0328DC5E08D18 CRC64;

Query Match Best Local Similarity 61.3%; Score 73; DB 2; Length 1577;

Matches 14; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 4 HLSTLEAWSNDNDTPYLHD 21
DB 551 HLSTLEAWSNDNDPPYXND 566

RESULT 10

O9ZAR4 PRELIMINARY; PRT; 1527 AA.

AC O9ZAR4; (TREMBlrel. 10, Created)

DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)

DE 01-OCT-2002 (TREMBlrel. 22, Last annotation update)

OS Leuconostoc mesenteroides.

OC Bacteria; Firmicutes; Lactobacillales; Leuconostoc.

OX NCBI_TaxID=1245;

RN SEQUENCE FROM N.A.

RA Bhatnagar R., Singh D.K.S.;

RT "Cloning and Molecular Characterization of Dextranucrase Gene from

Leuconostoc mesenteroides NRRL B-512^F."

DR EMBL; U81374; AAD10952.1; -

DR InterPro; IPR002479; CW binding.

DR Pfam; PF01473; CW binding_1; 16.

DR Pfam; PF02324; Glyco_hydro_70; 1.

SQ SEQUENCE 1527 AA; 169709 MW; IDFAFA237C743398 CRC64;

Query Match Best Local Similarity 59.7%; Score 71; DB 2; Length 1527;

Matches 14; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

OY 1 ANDHLSTLEAWSNDNDTPYLHD 21
DB 581 ANOHLSTLEAWSNDNDPLYVTD 601

RESULT 11

O8KRE1 PRELIMINARY; PRT; 1527 AA.

AC O8KRE1; (TREMBlrel. 22, Created)

DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)

DE 01-MAR-2003 (TREMBlrel. 23, Last annotation update)

OS Leuconostoc mesenteroides.

OC Bacteria; Firmicutes; Lactobacillales; Leuconostoc.

OX NCBI_TaxID=1245;

RN SEQUENCE FROM N.A.

RA Neubauer H., Bauche A., Mollet B.;

RT "Isolation and characterization of the dextranucrase Dsd of

Leuconostoc mesenteroides loc4."

Best Local Similarity 66.7%; Pred. No. 0.037;

Matches 14; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

OY 1 ANDHLSTLEAWSNDNDTPYLHD 21
DB 581 ANOHLSTLEAWSNDNDPLYVTD 601

RESULT 12

O8G9Q2 PRELIMINARY; PRT; 2835 AA.

AC O8G9Q2; (TREMBlrel. 23, Created)

DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)

DE 01-MAR-2003 (TREMBlrel. 23, Last annotation update)

OS Leuconostoc mesenteroides.

OC Bacteria; Firmicutes; Lactobacillales; Leuconostoc.

OX NCBI_TaxID=1245;

RN SEQUENCE FROM N.A.

RA MEDLINE=2231661; PubMed=12270834;

RA Boronnet S., Dolis-Lafargue M., Fabre E., Pizut S., Remaud-Simeon M.,

Mordan P., Willems R.M.;

RT "Molecular Characterisation of DSR-E, an alpha-1,2 linkage

synthesising dextranucrase with two catalytic domains."

DR EMBL; AJ430204; CAD22883.1; -

DR Transferase; Glycosyltransferase.

FT NON TER 1

SQ SEQUENCE 2835 AA; 313264 MW; D03262CDD735399D CRC64;

Query Match Best Local Similarity 58.0%; Score 69; DB 2; Length 2835;

Matches 11; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

OY 1 ANDHLSTLEAWSNDNDTPYL 19
DB 557 ANNHISLLEWDNDNDSDAYT 575

RESULT 13

O88542 PRELIMINARY; PRT; 1449 AA.

AC O88542; (TREMBlrel. 07, Created)

DT 01-AUG-1998 (TREMBlrel. 07, Last sequence update)

DE 01-AUG-1998 (TREMBlrel. 07, Last annotation update)

OS Streptococcus salivarius.

OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;

OX NCBI_TaxID=1304;

RN SEQUENCE FROM N.A.

RA Streptococcus salivarius VI477 gtfN."

RT Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.

DR EMBL; AF049609; AAC05156.1; -

DR InterPro; IPR002479; CW binding.

Best Local Similarity 72.2%; Pred. No. 0.21;

Matches 13; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 1 ANDHLSTLEAWSNDNDTPYLHD 21
DB 581 ANOHLSTLEAWSNDNDPLYVTD 601

RESULT 14

O8G9Q2 PRELIMINARY; PRT; 2835 AA.

AC O8G9Q2; (TREMBlrel. 23, Created)

DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)

DE 01-MAR-2003 (TREMBlrel. 23, Last annotation update)

OS Leuconostoc mesenteroides.

OC Bacteria; Firmicutes; Lactobacillales; Leuconostoc.

OX NCBI_TaxID=1245;

RN SEQUENCE FROM N.A.

RA MEDLINE=2231661; PubMed=12270834;

RA Boronnet S., Dolis-Lafargue M., Fabre E., Pizut S., Remaud-Simeon M.,

Mordan P., Willems R.M.;

RT "Molecular Characterisation of DSR-E, an alpha-1,2 linkage

synthesising dextranucrase with two catalytic domains."

DR EMBL; AJ430204; CAD22883.1; -

DR Transferase; Glycosyltransferase.

FT NON TER 1

SQ SEQUENCE 2835 AA; 313264 MW; D03262CDD735399D CRC64;

Query Match Best Local Similarity 55.5%; Score 66; DB 2; Length 1449;

Matches 13; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 1 ANDHLSTLEAWSNDNDTPYL 19
DB 557 ANNHISLLEWDNDNDSDAYT 575

RESULT 15

O88542 PRELIMINARY; PRT; 1449 AA.

AC O88542; (TREMBlrel. 07, Created)

DT 01-AUG-1998 (TREMBlrel. 07, Last sequence update)

DE 01-AUG-1998 (TREMBlrel. 07, Last annotation update)

OS Streptococcus salivarius.

OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;

OX NCBI_TaxID=1304;

RN SEQUENCE FROM N.A.

RA Streptococcus salivarius VI477 gtfN."

RT Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.

DR EMBL; AF049609; AAC05156.1; -

DR InterPro; IPR002479; CW binding.

QY 4 HLSTLEAMSDNDTPYLHD 21
 |||||
 DB 539 HLSTLEAMSHNDAYNED 556

RESULT 14
 ID Q55264 PRELIMINARY; PRT; 1449 AA.

AC Q55264;
 DT 01-NOV-1996 (TREMBlrel. 01, Created)
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
 DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
 DE Glucosyltransferase precursor.
 GN GTFU.
 OS Streptococcus salivarius.
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 OX NCBI_TaxID=1304;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=95122197; PubMed=7822030;
 RA Simpson C.L., Giffard P.M., Jacques N.A.;
 RT "Streptococcus salivarius ATCC 25975 possesses at least two genes
 coding for primer-independent glucosyltransferases."
 RL Infect. Immun. 63:609-621(1995).
 DR EMBL: L35495; AAC41412.1; -
 DR InterPro: IPR002479; CW_binding.
 DR InterPro: IPR003318; Glyco_hydro_70.
 DR Pfam: PF01473; CW_binding_1; 8.
 DR Pfam: PF02324; Glyco_hydro_70; 1.
 KW Signal; Transferase.
 FT SIGNAL 1 35 POTENTIAL.
 FT CHAIN 36 1449 GLUCOSYLTRANSFERASE.
 SQ SEQUENCE 1449 AA; 159984 MM; D622F07306B86A46 CRC64;

Query Match 55.5%; Score 66; DB 2; Length 1449;
 Best Local Similarity 72.2%; Pred. No. 0.21;
 Matches 13; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 4 HLSTLEAMSDNDTPYLHD 21
 |||||
 DB 539 HLSTLEAMSHNDAYNED 556

RESULT 15
 ID Q8KZL5 PRELIMINARY; PRT; 1554 AA.

AC Q8KZL5;
 DT 01-OCT-2002 (TREMBlrel. 22, Created)
 DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)
 DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
 DE Glucosyltransferase.
 GN GTFU.
 OS Streptococcus sobrinus.
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 OX NCBI_TaxID=1310;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21958684; PubMed=11960691;
 RA Hanada N., Fukushima K., Nomura Y., Sempuku H., Hayakawa M.,
 Mukasa H., Shiroza T., Abiko Y.;
 RT "Cloning and nucleotide sequence analysis of the Streptococcus
 sobrinus gtfu gene that produces a highly branched water-soluble
 glucan."
 RT Biochim. Biophys. Acta 1570:75-79(2002).
 DR EMBL: AB089438; BAC07265.1; -
 DR InterPro: IPR002479; CW_binding.
 DR InterPro: IPR003318; Glyco_hydro_70.
 DR Pfam: PF01473; CW_binding_1; 6.
 DR Pfam: PF02324; Glyco_hydro_70; 1.
 KW Transferase.

SEQ SEQUENCE 1554 AA; 171676 MM; 6981BCC1DAE24A73 CRC64;

Query Match 52.9%; Score 63; DB 2; Length 1554;
 Best Local Similarity 66.7%; Pred. No. 0.69;
 Matches 12; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 4 HLSTLEAMSDNDTPYLHD 21
 |||||
 DB 488 HLSTLEAMSLNDNQYNE 505

Search completed: November 13, 2003, 09:44:01
 UOB time : 31.6493 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 13, 2003, 09:31:40 ; Search time 13.6351 Seconds
(without alignments)
65.165 Million cell updates/sec

Title: US-09-290-049a-10
Perfect score: 119
Sequence: 1 ANDHSLLEWSDNDPTPLHD 21

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: /cgn2_6/ptodata/2/1aa/5A COMB.pep.*
2: /cgn2_6/ptodata/2/1aa/5B COMB.pep.*
3: /cgn2_6/ptodata/2/1aa/6A COMB.pep.*
4: /cgn2_6/ptodata/2/1aa/6B COMB.pep.*
5: /cgn2_6/ptodata/2/1aa/PTUS COMB.pep.*
6: /cgn2_6/ptodata/2/1aa/backfile1.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	119	100.0	1475 3 US-09-007-999-2	Sequence 2, Appl1
2	119	100.0	1475 3 US-09-210-361-2	Sequence 2, Appl1
3	119	100.0	1475 3 US-09-740-274-2	Sequence 2, Appl1
4	110	92.4	1375 4 US-09-210-361-4	Sequence 4, Appl1
5	110	92.4	1375 4 US-09-740-274-4	Sequence 4, Appl1
6	75	63.0	545 4 US-09-604-957-4	Sequence 2, Appl1
7	75	63.0	1430 3 US-09-008-172-2	Sequence 2, Appl1
8	75	63.0	1430 3 US-09-210-361-6	Sequence 6, Appl1
9	75	63.0	1430 4 US-09-740-274-6	Sequence 6, Appl1
10	71	59.7	523 4 US-09-604-957-5	Sequence 5, Appl1
11	63	52.9	1577 2 US-08-793-824-2	Sequence 2, Appl1
12	60	50.4	584 4 US-09-604-957-6	Sequence 6, Appl1
13	60	50.4	2057 4 US-09-499-203-2	Sequence 2, Appl1
14	49	41.2	535 4 US-09-604-957-7	Sequence 4, Appl1
15	49	41.2	1278 4 US-09-604-957-3	Sequence 3, Appl1
16	46	38.7	336 4 US-09-198-452A-662	Sequence 662, App
17	46	38.7	596 4 US-09-252-991A-21255	Sequence 2155, A
18	45	37.8	386 4 US-09-154-750A-80	Sequence 80, Appl
19	45	37.8	385 1 US-08-361-920-23	Sequence 23, Appl
20	45	37.8	385 1 US-08-479-939-23	Sequence 23, Appl
21	45	37.8	385 1 US-08-483-432-23	Sequence 23, Appl
22	44.5	37.4	130 4 US-09-328-352-4952	Sequence 4952, Ap
23	44.5	37.4	236 4 US-09-252-991A-21380	Sequence 21380, A
24	44	37.0	234 4 US-09-134-001C-4514	Sequence 4514, Ap
25	44	37.0	647 3 US-09-311-170-2	Sequence 2, Appl1
26	44	37.0	647 3 US-09-134-001C-5458	Sequence 5458, Ap
27	44	37.0	749 4 US-09-562-737-96	Sequence 96, Appl

ALIGNMENTS

28	43	36.1	2465 2 US-08-596-291-3	Sequence 3, Appl1
29	43	36.1	2465 3 US-09-100-804-3	Sequence 3, Appl1
30	43	36.1	2466 3 US-09-080-855-12	Sequence 12, Appl
31	43	36.1	2466 4 US-09-566-076-12	Sequence 12, Appl
32	43	36.1	2466 5 PCT-US94-09943-2	Sequence 46, Appl1
33	43	36.1	2485 5 US-09-290-640-46	Sequence 5029, Ap
34	42	35.3	211 4 US-09-107-532A-5029	Sequence 3, Appl1
35	42	35.3	259 1 US-08-277-231A-3	Sequence 6, Appl1
36	42	35.3	259 2 US-08-473-750-6	Sequence 6, Appl1
37	42	35.3	259 2 US-08-477-326-6	Sequence 23168, A
38	42	35.3	288 4 US-09-252-991A-23168	Sequence 16, Appl
39	42	35.3	989 2 US-08-070-301-16	Sequence 16731, A
40	42	35.3	1428 4 US-09-252-991A-30731	Sequence 28000, A
41	41.5	34.9	454 4 US-09-252-991A-28000	Sequence 10, Appl
42	41.5	34.9	1048 4 US-09-171-699-10	Sequence 22373, A
43	41	34.5	244 4 US-09-252-991A-22373	Sequence 38, Appl
44	41	34.5	270 4 US-09-323-872A-38	Sequence 7769, Ap
45	41	34.5	296 4 US-09-328-352-7769	

RESULT 1
US-09-007-999-2
; Sequence 2, Application US/09007999
; Patent No. 6087559
; GENERAL INFORMATION:
; APPLICANT: Nichols, Scott E.
; TITLE OF INVENTION: Substitutes for Modified starch and
; FILE REFERENCE: 0356D
; CURRENT APPLICATION NUMBER: US/09/007,999
; CURRENT FILING DATE: 1998-01-16
; EARLIER APPLICATION NUMBER: 08/478,704
; EARLIER FILING DATE: 1995-06-07
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 1475
; TYPE: PRT
; ORGANISM: Streptococcus mutans
US-09-007-999-2

Query Match 100.0%; Score 119; DB 3; Length 1475;
Best Local Similarity 100.0%; Pred. No. 1e-10;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ANDHSLLEWSDNDPTPLHD 21
Db 481 ANDHSLLEWSDNDPTPLHD 501
RESULT 2
US-09-210-361-2
; Sequence 2, Application US/09210361
; Patent No. 6284479
; GENERAL INFORMATION:
; APPLICANT: Nichols, Scott E.
; TITLE OF INVENTION: Substitutes for Modified Starches and
; FILE REFERENCE: 0357CR
; CURRENT APPLICATION NUMBER: US/09/210,361
; CURRENT FILING DATE: 1998-12-11
; EARLIER APPLICATION NUMBER: 09/007,999
; EARLIER FILING DATE: 1998-01-16
; EARLIER APPLICATION NUMBER: 08/478,704
; EARLIER FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 09/009,620
; EARLIER FILING DATE: 1998-01-20
; EARLIER APPLICATION NUMBER: 08/485,243
; EARLIER FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 09/008,172

```

; EARLIER FILING DATE: 1998-01-16
; EARLIER APPLICATION NUMBER: 08/482,711
; EARLIER FILING DATE: 1995-06-07
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 1475
; TYPE: PRT
; ORGANISM: Streptococcus mutans
US-09-210-361-2

Query Match          100.0%; Score 119; DB 3; Length 1475;
Best Local Similarity 100.0%; Pred. No. 1e-10;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 ANDHLSILEAWSNDTPYLHD 21
Db      481 ANDHLSILEAWSNDTPYLHD 501

RESULT 3
US-09-740-274-2
; Sequence 2, Application US/09740274
; Patent No. 6465203
; GENERAL INFORMATION:
; APPLICANT: Nichols, Scott E.
; TITLE OF INVENTION: Glucan-containing Compositions and Paper
; FILE REFERENCE: 0357CRD
; CURRENT APPLICATION NUMBER: US/09/740,274
; PRIOR FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: 09/210,361
; PRIOR FILING DATE: 1998-12-11
; PRIOR APPLICATION NUMBER: 09/007,999
; PRIOR FILING DATE: 1998-01-16
; PRIOR APPLICATION NUMBER: 08/478,704
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: 09/009,620
; PRIOR FILING DATE: 1998-01-20
; PRIOR APPLICATION NUMBER: 08/485,243
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: 09/008,172
; PRIOR FILING DATE: 1998-01-16
; PRIOR APPLICATION NUMBER: 08/482,711
; PRIOR FILING DATE: 1995-06-07
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 1475
; TYPE: PRT
; ORGANISM: Streptococcus mutans
US-09-740-274-2

Query Match          100.0%; Score 119; DB 4; Length 1475;
Best Local Similarity 100.0%; Pred. No. 1e-10;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 ANDHLSILEAWSNDTPYLHD 21
Db      481 ANDHLSILEAWSNDTPYLHD 501

RESULT 4
US-09-210-361-4
; Sequence 4, Application US/09210361
; Patent No. 6284479
; GENERAL INFORMATION:
; APPLICANT: Nichols, Scott E.
; TITLE OF INVENTION: Substitutes for Modified Starches and
; FILE REFERENCE: 0357CR
; CURRENT APPLICATION NUMBER: US/09/210,361
; PRIOR FILING DATE: 1998-12-11
; EARLIER APPLICATION NUMBER: 09/007,999
```

```

; EARLIER FILING DATE: 1998-01-16
; EARLIER APPLICATION NUMBER: 08/478,704
; EARLIER FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 09/009,620
; EARLIER FILING DATE: 1998-01-20
; EARLIER APPLICATION NUMBER: 08/485,243
; EARLIER FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 09/008,172
; EARLIER FILING DATE: 1998-01-16
; EARLIER APPLICATION NUMBER: 08/482,711
; EARLIER FILING DATE: 1995-06-07
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 1375
; TYPE: PRT
; ORGANISM: streptococcus mutans
US-09-740-274-4

Query Match          92.4%; Score 110; DB 3; Length 1375;
Best Local Similarity 95.2%; Pred. No. 2.8e-09;
Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 ANDHLSILEAWSNDTPYLHD 21
Db      507 ANDHLSILEAWSNDTPYLHD 527

RESULT 5
US-09-740-274-4
; Sequence 4, Application US/09740274
; Patent No. 6465203
; GENERAL INFORMATION:
; APPLICANT: Nichols, Scott E.
; TITLE OF INVENTION: Glucan-containing Compositions and Paper
; FILE REFERENCE: 0357CRD
; CURRENT APPLICATION NUMBER: US/09/740,274
; PRIOR FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: 09/210,361
; PRIOR FILING DATE: 1998-12-11
; PRIOR APPLICATION NUMBER: 09/007,999
; PRIOR FILING DATE: 1998-01-16
; PRIOR APPLICATION NUMBER: 08/478,704
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: 09/009,620
; PRIOR FILING DATE: 1998-01-20
; PRIOR APPLICATION NUMBER: 08/485,243
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: 09/008,172
; PRIOR FILING DATE: 1998-01-16
; PRIOR APPLICATION NUMBER: 08/482,711
; PRIOR FILING DATE: 1995-06-07
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 1375
; TYPE: PRT
; ORGANISM: streptococcus mutans
US-09-740-274-4

Query Match          92.4%; Score 110; DB 4; Length 1375;
Best Local Similarity 95.2%; Pred. No. 2.8e-09;
Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 ANDHLSILEAWSNDTPYLHD 21
Db      507 ANDHLSILEAWSNDTPYLHD 527

RESULT 6
US-09-604-957-4
; Sequence 4, Application US/09604957
; Patent No. 6486314
```



```

; GENERAL INFORMATION:
; APPLICANT: VAN GEEL-SCHUTTEN, GERRITDINA HENDRIKA
; APPLICANT: DIJKHUIZEN, LOBBERT
; APPLICANT: RAHMOUI, HAKIM
; APPLICANT: LEER, ROBERT-JAN
; TITLE OF INVENTION: NOVEL GLUCOSYLTRANSFERASE AND NOVEL GLUCAN
; FILE REFERENCE: BO 43388
; CURRENT APPLICATION NUMBER: US/09/604,957
; CURRENT FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: 00201871.1
; PRIOR FILING DATE: 2000-05-25
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 545
; TYPE: PRF
; ORGANISM: Streptococcus mutans
; US-09-604-957-4

Query Match      63.0%; Score 75; DB 4; Length 545;
Best Local Similarity 71.4%; Pred. No. 0.00043;
Matches 15; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY      1 ANDHSLTEAMSDNDPTYLHD 21
DB      75 AINHLSTLEAMSDNDPQYNKD 95

RESULT 7
US-09-008-172-2
; Sequence 2, Application US/09008172
; Patent No. 6127602
; GENERAL INFORMATION:
; APPLICANT: Nichols, Scott E.
; TITLE OF INVENTION: Substitutes for Modified Starch and
; FILE REFERENCE: 0358D
; CURRENT APPLICATION NUMBER: US/09/008,172
; CURRENT FILING DATE: 1998-01-16
; EARLIER APPLICATION NUMBER: 08/482,711
; EARLIER FILING DATE: 1995-06-07
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 1430
; TYPE: PRF
; ORGANISM: Streptococcus mutans
; US-09-008-172-2

Query Match      63.0%; Score 75; DB 3; Length 1430;
Best Local Similarity 71.4%; Pred. No. 0.0014;
Matches 15; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY      1 ANDHSLTEAMSDNDPTYLHD 21
DB      495 AINHLSTLEAMSDNDPQYNKD 515

RESULT 8
US-09-210-361-6
; Sequence 6, Application US/09210361
; Patent No. 6284479
; GENERAL INFORMATION:
; APPLICANT: Nichols, Scott E.
; TITLE OF INVENTION: Substitutes for Modified Starches and
; FILE REFERENCE: 0357CR
; CURRENT APPLICATION NUMBER: US/09/210,361
; CURRENT FILING DATE: 1998-12-11
; EARLIER APPLICATION NUMBER: 09/007,999
; EARLIER FILING DATE: 1998-01-16
; EARLIER APPLICATION NUMBER: 08/478,704
; EARLIER FILING DATE: 1995-06-07
```

```

; EARLIER APPLICATION NUMBER: 09/009,620
; EARLIER FILING DATE: 1998-01-20
; EARLIER APPLICATION NUMBER: 08/485,243
; EARLIER FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 09/008,172
; EARLIER FILING DATE: 1998-01-16
; EARLIER APPLICATION NUMBER: 08/482,711
; EARLIER FILING DATE: 1995-06-07
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 6
; LENGTH: 1430
; TYPE: PRF
; ORGANISM: streptococcus mutans
; US-09-210-361-6

Query Match      63.0%; Score 75; DB 3; Length 1430;
Best Local Similarity 71.4%; Pred. No. 0.0014;
Matches 15; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY      1 ANDHSLTEAMSDNDPTYLHD 21
DB      495 AINHLSTLEAMSDNDPQYNKD 515

RESULT 9
US-09-740-274-6
; Sequence 6, Application US/09740274
; Patent No. 6465203
; GENERAL INFORMATION:
; APPLICANT: Nichols, Scott E.
; TITLE OF INVENTION: Glucan-containing Compositions and Paper
; FILE REFERENCE: 0357CRD
; CURRENT APPLICATION NUMBER: US/09/740,274
; CURRENT FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: 09/210,361
; PRIOR FILING DATE: 1998-12-11
; PRIOR APPLICATION NUMBER: 09/007,999
; PRIOR FILING DATE: 1998-01-16
; PRIOR APPLICATION NUMBER: 08/478,704
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: 09/009,620
; PRIOR FILING DATE: 1998-01-20
; PRIOR APPLICATION NUMBER: 08/485,243
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: 09/008,172
; PRIOR FILING DATE: 1998-01-16
; PRIOR APPLICATION NUMBER: 08/482,711
; PRIOR FILING DATE: 1995-06-07
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 6
; LENGTH: 1430
; TYPE: PRF
; ORGANISM: streptococcus mutans
; US-09-740-274-6

Query Match      63.0%; Score 75; DB 4; Length 1430;
Best Local Similarity 71.4%; Pred. No. 0.0014;
Matches 15; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY      1 ANDHSLTEAMSDNDPTYLHD 21
DB      495 AINHLSTLEAMSDNDPQYNKD 515

RESULT 10
US-09-604-957-5
; Sequence 5, Application US/09604957
; Patent No. 6486314
; GENERAL INFORMATION:
; APPLICANT: VAN GEEL-SCHUTTEN, GERRITDINA HENDRIKA
; APPLICANT: DIJKHUIZEN, LOBBERT
```

APPLICANT: RAHAOUI, HAKIM
APPLICANT: LEBER, ROBERT-JAN
TITLE OF INVENTION: NOVEL GLUCOSYLTRANSFERASE AND NOVEL GLUCAN
FILE REFERENCE: BO 43388
CURRENT APPLICATION NUMBER: US/09/604,957
CURRENT FILING DATE: 2000-06-28
PRIOR APPLICATION NUMBER: 00201871.1
PRIOR FILING DATE: 2000-05-25
NUMBER OF SEQ ID NOS: 17
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 5
LENGTH: 523
TYPE: PRT
ORGANISM: Leuconostoc mesenteroides
US-09-604-957-5

Query Match 59.7%; Score 71; DB 4; Length 523;
Best Local Similarity 66.7%; Pred. No. 0.0018;
Matches 14; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 1 ANDHSTLEAMSNDNPTYLHD 21
Db 75 ANKHSTLEDMNGNDPQYVD 95

RESULT 11
US-08-793-824-2
Sequence 2, Application US/08793824
Patent No. 5981838
GENERAL INFORMATION:
APPLICANT: Simpson, Christine Lynn
APPLICANT: Giffard, Philip Morrison
APPLICANT: Jacques, Nicholas Anthony
TITLE OF INVENTION: Genetic Manipulation of Plants to
TITLE OF INVENTION: Increase Stored Carbohydrates
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: Griffith Hack & Co
STREET: Level 8, 168 Walker Street
CITY: No. 5981838th Sydney
STATE: New South Wales
COUNTRY: Australia
ZIP: 2060
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: IBM PC compatible
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/793,824
FILING DATE:
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: AT PM7643
FILING DATE: 24-AUG-1994
TELECOMMUNICATION INFORMATION:
TELEPHONE: 61 2 9957 5944
TELEFAX: 61 2 957 6288
TELEX: 26547
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1577 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: not relevant
MOLECULE TYPE: protein
ORIGINAL SOURCE:
ORGANISM: Streptococcus salivarius
US-08-793-824-2

Query Match 52.9%; Score 63; DB 2; Length 1577;
Best Local Similarity 72.2%; Pred. No. 0.14;
Matches 13; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 4 HSTLEAMSNDNPTYLHD 21
Db 591 HSTLEAMSNDHQNKD 608

RESULT 12
US-09-604-957-6
Sequence 6, Application US/09604957
Patent No. 6486314
GENERAL INFORMATION:
APPLICANT: VAN GEEL-SCHUTTEN, GERRITDINA HENDRIKA
APPLICANT: DIJKUIZEN, LUBBERT
APPLICANT: RAHAOUI, HAKIM
APPLICANT: LEBER, ROBERT-JAN
TITLE OF INVENTION: NOVEL GLUCOSYLTRANSFERASE AND NOVEL GLUCAN
FILE REFERENCE: BO 43388
CURRENT APPLICATION NUMBER: US/09/604,957
CURRENT FILING DATE: 2000-06-28
PRIOR APPLICATION NUMBER: 00201871.1
PRIOR FILING DATE: 2000-05-25
NUMBER OF SEQ ID NOS: 17
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 6
LENGTH: 584
TYPE: PRT
ORGANISM: Leuconostoc mesenteroides
US-09-604-957-6

Query Match 50.4%; Score 60; DB 4; Length 584;
Best Local Similarity 55.0%; Pred. No. 0.13;
Matches 11; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 1 ANDHSTLEAMSNDNPTYLH 20
Db 75 ANKHSTLEDMNGNDPQYVN 94

RESULT 13
US-09-499-203-2
Sequence 2, Application US/09499203
Patent No. 6570065
GENERAL INFORMATION:
APPLICANT: KOSGMANN, Jens
APPLICANT: WELSH, Thomas
APPLICANT: KUNZ, Martin
APPLICANT: KUNZ, Karola
TITLE OF INVENTION: Nucleic Acid Molecules Encoding Alternansucrase
FILE REFERENCE: 147-196P
CURRENT APPLICATION NUMBER: US/09/499,203
CURRENT FILING DATE: 2000-02-08
NUMBER OF SEQ ID NOS: 54
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 2
LENGTH: 2057
TYPE: PRT
ORGANISM: Leuconostoc mesenteroides
US-09-499-203-2

Query Match 50.4%; Score 60; DB 4; Length 2057;
Best Local Similarity 55.0%; Pred. No. 0.57;
Matches 11; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 1 ANDHSTLEAMSNDNPTYLH 20
Db 665 ANKHSTLEDMNGKDPQYVN 684

RESULT 14
US-09-604-957-7
Sequence 7, Application US/09604957
Patent No. 6486314
GENERAL INFORMATION:

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 13, 2003, 09:45:40 ; Search time 27.4692 Seconds
(without alignments)
139.565 Million cell updates/sec

Title: US-09-290-049A-10
Perfect score: 119
Sequence: 1 ANDHSLTEAWSDNDPTLYHD 21

Scoring table:
BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 666188 seqs, 182559486 residues

Total number of hits satisfying chosen parameters: 666188

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Published Applications AA:*
1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep:*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep:*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	119	100.0	1475	9 US-09-740-274-2	Sequence 2, Appl1
2	110	92.4	1375	9 US-09-740-274-4	Sequence 4, Appl1
3	75	63.0	545	10 US-09-995-749A-10	Sequence 10, Appl1
4	75	63.0	1430	9 US-09-740-274-6	Sequence 6, Appl1
5	71	59.7	522	10 US-09-995-749A-11	Sequence 11, Appl1
6	60	50.4	584	10 US-09-995-749A-12	Sequence 12, Appl1
7	51	42.9	502	12 US-10-009-823A-1	Sequence 1, Appl1
8	49	41.2	302	9 US-09-815-242-14004	Sequence 14004, A
9	49	41.2	535	10 US-09-995-749A-13	Sequence 13, Appl1
10	49	41.2	1781	10 US-09-995-749A-2	Sequence 2, Appl1
11	47.5	39.9	534	10 US-09-344-882-24	Sequence 24, Appl1
12	47.5	39.9	534	15 US-10-293-865-24	Sequence 24, Appl1
13	47	39.5	15	12 US-10-295-693-90	Sequence 90, Appl1
14	47	39.5	15	15 US-10-067-649-90	Sequence 90, Appl1
15	46.5	39.1	1604	9 US-09-888-615-73	Sequence 73, Appl1

16	46	38.7	199	10 US-09-738-626-4571	Sequence 4571, Ap
17	46	38.7	476	11 US-09-908-229-9	Sequence 9, Appl1
18	46	38.7	539	15 US-10-230-026-34	Sequence 34, Appl1
19	45.5	38.2	156	12 US-10-001-245-125	Sequence 125, App
20	45	37.8	268	15 US-10-156-761-13823	Sequence 167, App
21	44.5	37.4	210	15 US-10-127-032-167	Sequence 9158, Ap
22	44	37.0	216	15 US-10-156-761-9158	Sequence 11634, A
23	44	37.0	332	9 US-09-815-242-11634	Sequence 10188, A
24	44	37.0	331	9 US-09-815-242-10188	Sequence 13981, A
25	44	37.0	334	9 US-09-815-242-13981	Sequence 96, Appl1
26	44	37.0	749	15 US-10-211-962-96	Sequence 7, Appl1
27	44	37.0	816	15 US-10-080-114A-7	Sequence 126, App
28	43.5	36.6	155	12 US-10-001-245-126	Sequence 3308, Ap
29	43	36.1	41	11 US-09-764-891-3308	Sequence 18, Appl1
30	43	36.1	311	11 US-09-878-781-18	Sequence 9, Appl1
31	43	36.1	358	11 US-09-095-478-9	Sequence 7, Appl1
32	43	36.1	381	11 US-09-095-478-7	Sequence 7562, Ap
33	43	36.1	715	15 US-10-156-761-7562	Sequence 56, Appl1
34	43	36.1	1267	15 US-10-059-585-56	Sequence 12, Appl1
35	43	36.1	2466	12 US-10-177-980-12	Sequence 46, Appl1
36	43	36.1	2485	9 US-09-802-669-46	Sequence 2, Appl1
37	43	36.1	2654	15 US-10-227-610-2	Sequence 10190, A
38	42.5	35.7	491	9 US-09-815-242-10190	Sequence 13756, A
39	42.5	35.7	491	9 US-09-815-242-13756	Sequence 40, Appl1
40	42	35.3	31	15 US-10-092-908-40	Sequence 14, Appl1
41	42	35.3	41	15 US-10-092-908-41	Sequence 14, Appl1
42	42	35.3	95	9 US-09-945-301-14	Sequence 26, Appl1
43	42	35.3	298	11 US-09-988-626-298	Sequence 8655, Ap
44	42	35.3	348	10 US-09-738-626-348	Sequence 8339, Ap
45	42	35.3	484	15 US-10-156-761-8339	Sequence 8339, Ap

ALIGNMENTS

RESULT 1
US-09-740-274-2
; Sequence 2, Application US/09740274
; Patent No. US20020031826A1
GENERAL INFORMATION:
; APPLICANT: Nicholse, Scott E.
; TITLE OF INVENTION: Glucan-containing Compositions and Paper
; FILE REFERENCE: 03570RD
; CURRENT APPLICATION NUMBER: US/09/740, 274
; CURRENT FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: 09/210,361
; PRIOR FILING DATE: 1998-12-11
; PRIOR FILING DATE: 1998-12-11
; PRIOR APPLICATION NUMBER: 09/007, 999
; PRIOR FILING DATE: 1998-01-16
; PRIOR APPLICATION NUMBER: 08/478,704
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: 09/006,172
; PRIOR FILING DATE: 1998-01-16
; PRIOR APPLICATION NUMBER: 08/482,711
; PRIOR FILING DATE: 1995-06-07
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 1475
; TYPE: PRT
; ORGANISM: Streptococcus mutans
US-09-740-274-2

Query Match 100.0%; Score 119; DB 9; Length 1475;
Best Local Similarity 100.0%; Pred. No. 1.8e-09;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ANDHSLTEAWSDNDPTLYHD 21
|||||

Db 481 ANDHLSILEAWSNDPQYIHD 501

RESULT 2

US-09-740-274-4
; Sequence 4, Application US/09740274
; Patent No. US20020031826A1
; GENERAL INFORMATION:
; APPLICANT: Nichols, Scott E.
; TITLE OF INVENTION: Glucan-containing Compositions and Paper
; FILE REFERENCE: 0357CRD
; CURRENT APPLICATION NUMBER: US/09/740,274
; CURRENT FILING DATE: 2000-12-19
; PRIOR FILING DATE: 1998-12-11
; PRIOR APPLICATION NUMBER: 09/210,361
; PRIOR FILING DATE: 1998-12-11
; PRIOR APPLICATION NUMBER: 09/007,999
; PRIOR FILING DATE: 1998-01-16
; PRIOR APPLICATION NUMBER: 08/478,704
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: 09/009,620
; PRIOR FILING DATE: 1998-01-20
; PRIOR APPLICATION NUMBER: 08/485,243
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: 09/008,172
; PRIOR FILING DATE: 1998-01-16
; PRIOR APPLICATION NUMBER: 08/482,711
; PRIOR FILING DATE: 1995-06-07
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 1375
; TYPE: PRT
; ORGANISM: streptococcus mutans
US-09-740-274-4

Query Match 92.4%; Score 110; DB 9; Length 1375;

Best Local Similarity 95.2%; Pred. No. 4e-08; 1; Indels 0; Gaps 0;

Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ANDHLSILEAWSNDPQYIHD 21

Db 507 ANDHLSILEAWSNDPQYIHD 527

RESULT 3

US-09-995-749A-10
; Sequence 10, Application US/09995749A
; Patent No. US20020155568A1
; GENERAL INFORMATION:
; APPLICANT: VAN GEEL-SCHUTTEN, GERTRIDINA HENDRIKA
; APPLICANT: DIKHUIZEN, LOBBERT
; APPLICANT: RAHAOUTI, HAKIM
; APPLICANT: LEER, ROBERT-JAN
; TITLE OF INVENTION: NOVEL GLUCOSYLTRANSFERASES
; FILE REFERENCE: BO43388-CIP
; CURRENT APPLICATION NUMBER: US/09/995,749A
; CURRENT FILING DATE: 2001-11-29
; PRIOR FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: EPO 00201871.1
; PRIOR FILING DATE: 2000-05-25
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
; LENGTH: 545
; TYPE: PRT
; ORGANISM: Streptococcus mutans
US-09-995-749A-10

Query Match 63.0%; Score 75; DB 10; Length 545;

Best Local Similarity 71.4%; Pred. No. 0.0033; 5; Indels 0; Gaps 0;

Matches 15; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 ANDHLSILEAWSNDPQYIHD 21

Db 75 AINHLSILEAWSNDPQYIHD 95

RESULT 4

US-09-740-274-6
; Sequence 6, Application US/09740274
; Patent No. US20020031826A1
; GENERAL INFORMATION:
; APPLICANT: Nichols, Scott E.
; TITLE OF INVENTION: Glucan-containing Compositions and Paper
; FILE REFERENCE: 0357CRD
; CURRENT APPLICATION NUMBER: US/09/740,274
; CURRENT FILING DATE: 2000-12-19
; PRIOR FILING DATE: 1998-12-11
; PRIOR APPLICATION NUMBER: 09/210,361
; PRIOR FILING DATE: 1998-12-11
; PRIOR APPLICATION NUMBER: 09/007,999
; PRIOR FILING DATE: 1998-01-16
; PRIOR APPLICATION NUMBER: 08/478,704
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: 09/009,620
; PRIOR FILING DATE: 1998-01-20
; PRIOR APPLICATION NUMBER: 08/485,243
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: 09/008,172
; PRIOR FILING DATE: 1998-01-16
; PRIOR APPLICATION NUMBER: 08/482,711
; PRIOR FILING DATE: 1995-06-07
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 6
; LENGTH: 1430
; TYPE: PRT
; ORGANISM: streptococcus mutans
US-09-740-274-6

Query Match 63.0%; Score 75; DB 9; Length 1430;

Best Local Similarity 71.4%; Pred. No. 0.0097; 5; Indels 0; Gaps 0;

Matches 15; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 ANDHLSILEAWSNDPQYIHD 21

Db 495 AINHLSILEAWSNDPQYIHD 515

RESULT 5

US-09-995-749A-11
; Sequence 11, Application US/09995749A
; Patent No. US20020155568A1
; GENERAL INFORMATION:
; APPLICANT: VAN GEEL-SCHUTTEN, GERTRIDINA HENDRIKA
; APPLICANT: DIKHUIZEN, LOBBERT
; APPLICANT: RAHAOUTI, HAKIM
; APPLICANT: LEER, ROBERT-JAN
; TITLE OF INVENTION: NOVEL GLUCOSYLTRANSFERASES
; FILE REFERENCE: BO43388-CIP
; CURRENT APPLICATION NUMBER: US/09/995,749A
; CURRENT FILING DATE: 2001-11-29
; PRIOR FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: EPO 00201871.1
; PRIOR FILING DATE: 2000-05-25
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 11
; LENGTH: 522
; TYPE: PRT
; ORGANISM: Leuconostoc mesenteroides
US-09-995-749A-11

Query Match 59.7%; Score 71; DB 10; Length 522;

Best Local Similarity 66.7%; Pred. No. 0.013;

Matches 14; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 1 ANDHLSILEAMSDNDTPYLHD 21
Db 75 ANOHLSTLEDMWGNHNDPLYVTD 95

RESULT 6

US-09-995-749A-12
; Sequence 12, Application US/09995749A
; Patent No. US20020155568A1
; GENERAL INFORMATION:
; APPLICANT: VAN GEEL-SCHUTTEN, GERRITDINA HENDRIKA
; APPLICANT: DIJKHUIZEN, LUBBERT
; APPLICANT: RAHAOUI, HAKIM
; APPLICANT: LEER, ROBERT-JAN
; TITLE OF INVENTION: NOVEL GLUCOSYLTRANSFERASES
; FILE REFERENCE: B043388-CIP
; CURRENT APPLICATION NUMBER: US/09/995,749A
; CURRENT FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: 09/604,957
; PRIOR FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: EPO 00201871.1
; PRIOR FILING DATE: 2000-05-25
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 12
; LENGTH: 584
; TYPE: PRT
; ORGANISM: Leuconostoc mesenteroides
US-09-995-749A-12

Query Match 50.4%; Score 60; DB 10; Length 584;
Best Local Similarity 55.0%; Pred. No. 0.72;
Matches 11; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 1 ANDHLSILEAMSDNDTPYLHD 20
Db 75 ANKHLSTLEDMWGNHNDPLYVYV 94

RESULT 7

US-10-009-823A-1
; Sequence 1, Application US/10009823A
; Publication No. US20030157120A1
; GENERAL INFORMATION:
; APPLICANT: Fanuccio, Michael
; APPLICANT: Rosey, Everett Lee
; APPLICANT: Sinistraj, Meri
; APPLICANT: Hasse, Deleef
; APPLICANT: Parsons, Jim
; APPLICANT: Ankenbauer, Robert G.
; TITLE OF INVENTION: LAWSONIA DERIVED GENE AND RELATED FLGE
; FILE REFERENCE: POLYPEPTIDES, PEPTIDES AND PROTEINS AND THEIR USES
; CURRENT APPLICATION NUMBER: US/10/009,823A
; CURRENT FILING DATE: 2002-08-13
; PRIOR APPLICATION NUMBER: PCT/AU00/00437
; PRIOR FILING DATE: 2000-05-11
; PRIOR APPLICATION NUMBER: US 60/133,973
; PRIOR FILING DATE: 1999-05-13
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 502
; TYPE: PRT
; ORGANISM: Lawsonia intracellularis
US-10-009-823A-1

Query Match 42.9%; Score 51; DB 12; Length 502;
Best Local Similarity 47.1%; Pred. No. 14;
Matches 8; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 1 ANDHLSILEAMSDNDTP 17
Db 185 ANPYFALHESWKGNGTP 201

RESULT 8

US-09-815-242-14004
; Sequence 14004, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA-011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14004
; LENGTH: 302
; TYPE: PRT
; ORGANISM: Salmonella typhi
US-09-815-242-14004

Query Match 41.2%; Score 49; DB 9; Length 302;
Best Local Similarity 50.0%; Pred. No. 17;
Matches 10; Conservative 1; Mismatches 9; Indels 0; Gaps 0;

QY 2 NDHLSILEAMSDNDTPYLHD 21
Db 267 NDRLVIOEVLEASDIPYRHD 286

RESULT 9

US-09-995-749A-13
; Sequence 13, Application US/09995749A
; Patent No. US20020155568A1
; GENERAL INFORMATION:
; APPLICANT: VAN GEEL-SCHUTTEN, GERRITDINA HENDRIKA
; APPLICANT: DIJKHUIZEN, LUBBERT
; APPLICANT: RAHAOUI, HAKIM
; APPLICANT: LEER, ROBERT-JAN
; TITLE OF INVENTION: NOVEL GLUCOSYLTRANSFERASES
; FILE REFERENCE: B043388-CIP
; CURRENT APPLICATION NUMBER: US/09/995,749A
; CURRENT FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: 09/604,957
; PRIOR FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: EPO 00201871.1
; PRIOR FILING DATE: 2000-05-25
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 13
LENGTH: 535
TYPE: PRT
ORGANISM: Lactobacillus reuteri
US-09-995-749A-13

Query Match 41.2%; Score 49; DB 10; Length 535;
Best Local Similarity 40.0%; Pred. No. 31;
Matches 8; Conservative 5; Mismatches 7; Indels 0; Gaps 0;

QY 1 ANDHSLFAMSGNDTPYLH 20
DB 74 SNKHINLEDMWHADEPEYN 93

RESULT 10
US-09-995-749A-2
Sequence 2, Application US/09995749A
Patent No. US20020155568A1
GENERAL INFORMATION:
APPLICANT: VAN GEEL-SCHUTTEN, GERRITDINA HENDRIKA
APPLICANT: DIJKHUIZEN, LUBBERT
APPLICANT: RAHACUI, HAKIM
APPLICANT: LEBER, ROBERT-JAN
TITLE OF INVENTION: NOVEL GLUCOSYLTRANSFERASES
FILE REFERENCE: B043388-CIP
CURRENT APPLICATION NUMBER: US/09/995,749A
CURRENT FILING DATE: 2001-11-29
PRIOR APPLICATION NUMBER: 09/604,957
PRIOR FILING DATE: 2000-06-28
PRIOR APPLICATION NUMBER: EPO 00201871.1
PRIOR FILING DATE: 2000-05-25
NUMBER OF SEQ ID NOS: 19
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 2
LENGTH: 1781
TYPE: PRT
ORGANISM: Lactobacillus reuteri
US-09-995-749A-2

Query Match 41.2%; Score 49; DB 10; Length 1781;
Best Local Similarity 40.0%; Pred. No. 1,2e+02;
Matches 8; Conservative 5; Mismatches 7; Indels 0; Gaps 0;

QY 1 ANDHSLFAMSGNDTPYLH 20
DB 1053 SNKHINLEDMWHADEPEYN 1072

RESULT 11
US-09-344-882-24
Sequence 24, Application US/09344882
Patent No. US20020162137A1
GENERAL INFORMATION:
APPLICANT: Nikolaou, Basil J
APPLICANT: Wurttele, Eve S
APPLICANT: Oliver, David J
APPLICANT: Behal, Robert
APPLICANT: Schnable, Patrick S
APPLICANT: Ke, Jinshan
APPLICANT: Johnson, Jerry L
APPLICANT: Allied, Carolyn C
APPLICANT: Fattland, Beth
APPLICANT: Lutziger, Isabelle
APPLICANT: Wen, Tsui-Jung
TITLE OF INVENTION: Materials and Methods for the Alteration of Enzyme and
FILE REFERENCE: 201573
CURRENT APPLICATION NUMBER: US/09/344,882
CURRENT FILING DATE: 1999-06-25
PRIOR APPLICATION NUMBER: US 60/090,717
PRIOR FILING DATE: 1998-06-26
NUMBER OF SEQ ID NOS: 38

SOFTWARE: PatentIn Ver. 2.2
SEQ ID NO 24
LENGTH: 534
TYPE: PRT
ORGANISM: Arabidopsis thaliana
US-09-344-882-24

Query Match 39.9%; Score 47.5; DB 10; Length 534;
Best Local Similarity 52.9%; Pred. No. 53;
Matches 9; Conservative 2; Mismatches 5; Indels 1; Gaps 1;

QY 2 NDHSLFAMSGNDTPY 18
DB 132 NDEIALEFTW-DNGKPY 147

RESULT 12
US-10-293-865-24
Sequence 24, Application US/10293865
Publication No. US20030106090A1
GENERAL INFORMATION:
APPLICANT: Nikolaou, Basil J
APPLICANT: Wurttele, Eve S
APPLICANT: Oliver, David J
APPLICANT: Behal, Robert
APPLICANT: Schnable, Patrick S
APPLICANT: Ke, Jinshan
APPLICANT: Johnson, Jerry L
APPLICANT: Allied, Carolyn C
APPLICANT: Fattland, Beth
APPLICANT: Lutziger, Isabelle
APPLICANT: Wen, Tsui-Jung
TITLE OF INVENTION: Materials and Methods for the Alteration of Enzyme and
FILE REFERENCE: 217113
CURRENT APPLICATION NUMBER: US/10/293,865
CURRENT FILING DATE: 2002-11-13
PRIOR APPLICATION NUMBER: US 09/344,882
PRIOR FILING DATE: 1999-06-25
PRIOR APPLICATION NUMBER: US 60/090,717
PRIOR FILING DATE: 1998-06-26
NUMBER OF SEQ ID NOS: 38
SOFTWARE: PatentIn Ver. 3.1
SEQ ID NO 24
LENGTH: 534
TYPE: PRT
ORGANISM: Arabidopsis thaliana
US-10-293-865-24

Query Match 39.9%; Score 47.5; DB 15; Length 534;
Best Local Similarity 52.9%; Pred. No. 53;
Matches 9; Conservative 2; Mismatches 5; Indels 1; Gaps 1;

QY 2 NDHSLFAMSGNDTPY 18
DB 132 NDEIALEFTW-DNGKPY 147

RESULT 13
US-10-295-693-90
Sequence 90, Application US/10295693
Publication No. US20030198976A1
GENERAL INFORMATION:
APPLICANT: Bristol-Myers Squibb Company
APPLICANT: Briscoe, MYERS G-PROTEIN COUPLED RECEPTOR, HGPBBMT14, RELATED TO THE
TITLE OF INVENTION: GPCR, GPR73
FILE REFERENCE: D0118 CIP
CURRENT APPLICATION NUMBER: US/10/295,693
CURRENT FILING DATE: 2002-11-14
PRIOR APPLICATION NUMBER: US 60/266,525
PRIOR FILING DATE: 2001-02-05
PRIOR APPLICATION NUMBER: US 10/067,649
PRIOR FILING DATE: 2002-02-05

```

; PRIOR APPLICATION NUMBER: US 60/329,897
; PRIOR FILING DATE: 2001-10-16
; NUMBER OF SEQ ID NOS: 95
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 90
; LENGTH: 15
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthesized Random Peptide.
US-10-295-693-90

Query Match      39.5%; Score 47; DB 12; Length 15;
Best Local Similarity 61.5%; Pred. No. 1.2;
Matches 8; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY      8 LEAMSDNDTPYLH 20
Db      3 LEAMDLSDTPHLV 15

RESULT 14
US-10-067-649-90
; Sequence 90, Application US/10067649
; Publication No. US2003010057A1
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: A NOVEL HUMAN G-PROTEIN COUPLED RECEPTOR, HGPBRM14, RELATED TO T
; TITLE OF INVENTION: ORPHAN GPCR, GPR73
; FILE REFERENCE: D0118 NP
; CURRENT APPLICATION NUMBER: US/10/067,649
; CURRENT FILING DATE: 2002-02-05
; PRIOR APPLICATION NUMBER: US 60/266,525
; PRIOR FILING DATE: 2001-02-05
; PRIOR APPLICATION NUMBER: US 60/329,897
; PRIOR FILING DATE: 2001-10-16
; NUMBER OF SEQ ID NOS: 92
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 90
; LENGTH: 15
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthesized Random Peptide.
US-10-067-649-90

Query Match      39.5%; Score 47; DB 15; Length 15;
Best Local Similarity 61.5%; Pred. No. 1.2;
Matches 8; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY      8 LEAMSDNDTPYLH 20
Db      3 LEAMDLSDTPHLV 15

RESULT 15
US-09-888-615-73
; Sequence 73, Application US/09888615
; Patent No. US20020064856A1
; GENERAL INFORMATION:
; APPLICANT: PLOMAN, GREGORY
; APPLICANT: WHYTE, DAVID
; APPLICANT: CAENEPEL, SEAN
; APPLICANT: CHARYDCZAK, GLEN
; APPLICANT: MANNING, GERARD
; APPLICANT: SUDARSANAM, SUCHA
; TITLE OF INVENTION: NOVEL PROTEASES
; FILE REFERENCE: 038602/1214
; CURRENT APPLICATION NUMBER: US/09/888,615
; CURRENT FILING DATE: 2001-06-26
; PRIOR APPLICATION NUMBER: 60/214,047
; PRIOR FILING DATE: 2000-06-26
; NUMBER OF SEQ ID NOS: 150
```

```

; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 73
; LENGTH: 1604
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-888-615-73

Query Match      39.1%; Score 46.5; DB 9; Length 1604;
Best Local Similarity 47.4%; Pred. No. 2.6e+02;
Matches 9; Conservative 3; Mismatches 4; Indels 3; Gaps 1;

QY      5 LSLLEAMSDN---DTPYLH 20
Db      294 VALLEVMKDNRTDIPHLH 312

Search completed: November 13, 2003, 10:29:01
Job time : 27.4692 secs
```


PI Nichols SE;
 XX WPI; 2002-414332/44.
 DR N-PSDB; ABK52940.
 XX
 PT Glucosyltransferase B or D protein useful for producing a glucan useful
 as substitutes for and additions to modified starch and latexes in
 PT paper manufacture, comprises mutations in specific positions -
 XX
 PS Disclosure; Page 38-42; 44pp; English.

XX The invention an isolated protein comprising a glucosyltransferase
 CC (GTF) B polypeptide having changes at position from I448V, D457N,
 CC D567T, K1014T, D457N/D567T, D457N/D571K, D567T/D571K,
 CC D567T/D571K/K1014T, I448V/D457N/D567T/D571K/K779Q/K1014T,
 CC Y169A/Y170A/Y171A, and K779Q or a GTF D polypeptide having
 CC changes at positions from T589D, T589E, N471D, N471D/T589D, and
 CC N471D/T589E. Also included are a glucan produced by the GTF mutant,
 CC an isolated polynucleotide which encodes P1 or P2, or its complementary
 CC polynucleotide, a ribonucleic acid sequence encoding the GTF mutant,
 CC an expression cassette comprising the polynucleotide operably linked to a
 CC promoter, a vector comprising the expression cassette, host cell
 CC introduced with the vector, a transgenic plant comprising the
 CC vector, a seed or tuber from the transgenic plant, a paper sizing and/or
 CC coating composition comprising a glucan produced in a plant transformed
 CC with a gene encoding the mutant GTF, wild type or starch, a latex,
 CC thermoplastic molecule or their combinations or glucan and starch where
 CC the glucan is produced in the amyloplast and/or vacuole or a maize line
 CC deficient in starch biosynthesis, transformed with a gene encoding a
 CC glucosyltransferase B or D enzyme, wild-type or mutant and a paper
 CC comprising the glucan (paper sizing/coating agent). The vector is useful
 CC for producing a glucan in a plant. The method comprises transforming a
 CC plant cell with the vector, growing the plant cell under plant growing
 CC conditions to produce a regenerated plant and inducing expression of the
 CC polynucleotide for a time sufficient to produce the glucan in the
 CC regenerated plant, where the vector contains a transit sequence from
 CC ribulose biphosphate carboxylase small subunit, waxy, brittle-1 and
 CC chlorophyll AB binding protein to produce a transgenic plant, and glucan
 CC is produced in the amyloplast of potato or the vacuole of sugar beet.
 CC Glucans are useful as substitutes for and additions to modified starch
 CC and latexes in paper manufacture. Unlike prior art techniques, which
 CC require input materials that produce chemical effluents, paper
 CC manufacture utilizing the glucan produced by GTF, which utilizes
 CC biologically produced input materials, is more cost-effective and
 CC environmentally friendly. Moreover, glucans also exhibit thermoplastic
 CC properties and impart gloss to the paper during coating step.
 CC The present sequence represents GTFD.

XX Sequence 1430 AA;

XX Query Match 100.0%; Score 116; DB 23; Length 1430;
 Best Local Similarity 100.0%; Pred. No. 2.6e-09;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AINHLSTLEAWSDNDPOYKND 21
 ||||||||||||||||||||
 Db 495 AINHLSTLEAWSDNDPOYKND 515

RESULT 2
 AAU98041 standard; Protein; 1430 AA.
 ID AAU98041;
 AC AAU98041;
 XX
 DT 27-AUG-2002 (first entry)

XX S. mutans glucosyltransferase GTFD mutant T589D.
 DE
 XX Glucosyltransferase; GTFD; transgenic plant; paper sizing;
 KM coating composition; glucan; starch; latex; thermoplastic molecule;
 KM amyloplast; vacuole; paper manufacture; mutant; mutlein.

OS Streptococcus mutans.
 OS Synthetic.
 FH Key Location/Qualifiers
 FT Misc-difference 589 /note= "wild-type Thr substituted by Asp"
 XX US2002031826-A1.
 PN 14-MAR-2002.
 PD 19-DEC-2000; 2000US-0740274.

XX 11-DEC-1998; 98US-0210361.
 XX 07-JUN-1995; 95US-0478704.
 PR 07-JUN-1995; 95US-0482711.
 PR 07-JUN-1995; 95US-0485243.
 PR 16-JAN-1998; 98US-0007899.
 PR 16-JAN-1998; 98US-0008172.
 PR 20-JAN-1998; 98US-0009620.

XX (NICH/) NICHOLS S E.

XX Nichols SE;

XX WPI; 2002-414332/44.

XX Glucosyltransferase B or D protein useful for producing a glucan useful
 as substitutes for and additions to modified starch and latexes in
 PT paper manufacture, comprises mutations in specific positions -
 PT

XX Claim 36; Page -: 44pp; English.

XX The invention an isolated protein comprising a glucosyltransferase
 CC (GTF) B polypeptide having changes at position from I448V, D457N,
 CC D567T, K1014T, D457N/D567T, D457N/D571K, D567T/D571K,
 CC D567T/D571K/K1014T, I448V/D457N/D567T/D571K/K779Q/K1014T,
 CC Y169A/Y170A/Y171A, and K779Q or a GTF D polypeptide having
 CC changes at positions from T589D, T589E, N471D, N471D/T589D, and
 CC N471D/T589E. Also included are a glucan produced by the GTF mutant,
 CC an isolated polynucleotide which encodes P1 or P2, or its complementary
 CC polynucleotide, a ribonucleic acid sequence encoding the GTF mutant,
 CC an expression cassette comprising the polynucleotide operably linked to a
 CC promoter, a vector comprising the expression cassette, host cell
 CC introduced with the vector, a transgenic plant comprising the
 CC vector, a seed or tuber from the transgenic plant, a paper sizing and/or
 CC coating composition comprising a glucan produced in a plant transformed
 CC with a gene encoding the mutant GTF, wild type or starch, a latex,
 CC thermoplastic molecule or their combinations or glucan and starch where
 CC the glucan is produced in the amyloplast and/or vacuole or a maize line
 CC deficient in starch biosynthesis, transformed with a gene encoding a
 CC glucosyltransferase B or D enzyme, wild-type or mutant and a paper
 CC comprising the glucan (paper sizing/coating agent). The vector is useful
 CC for producing a glucan in a plant. The method comprises transforming a
 CC plant cell with the vector, growing the plant cell under plant growing
 CC conditions to produce a regenerated plant and inducing expression of the
 CC polynucleotide for a time sufficient to produce the glucan in the
 CC regenerated plant, where the vector contains a transit sequence from
 CC ribulose biphosphate carboxylase small subunit, waxy, brittle-1 and
 CC chlorophyll AB binding protein to produce a transgenic plant, and glucan
 CC is produced in the amyloplast of potato or the vacuole of sugar beet.
 CC Glucans are useful as substitutes for and additions to modified starch
 CC and latexes in paper manufacture. Unlike prior art techniques, which
 CC require input materials that produce chemical effluents, paper
 CC manufacture utilizing the glucan produced by GTF, which utilizes
 CC biologically produced input materials, is more cost-effective and
 CC environmentally friendly. Moreover, glucans also exhibit thermoplastic
 CC properties and impart gloss to the paper during coating step.
 CC The present sequence represents a GTFD mutant of the invention.
 CC Note: The present sequence is not shown in the specification but
 CC was created by the indexer using the GTFD sequence appearing as AAU98029
 XX and the information in claim 36.

SQ Sequence 1430 AA;

Query Match 100.0%; Score 116; DB 23; Length 1430;
 Best Local Similarity 100.0%; Pred. No. 2.6e-09;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 AINHLSTLEAWSNDPQYKND 21
 |||||
 495 AINHLSTLEAWSNDPQYKND 515

RESULT 3
 AAU98042
 ID AAU98042 standard; Protein; 1430 AA.
 XX AAU98042;
 AC AAU98042;
 XX 27-AUG-2002 (first entry)
 DT 27-AUG-2002 (first entry)
 DE S. mutans glucosyltransferase GTFD mutant T589E.
 XX Glucosyltransferase; GTFD; transgenic plant; paper sizing;
 KM coating composition; glucan; starch; latex; thermoplastic molecule;
 KW amyloplast; vacuole; paper manufacture; mutant; mutain.
 XX Streptococcus mutans.
 OS Synthetic.
 XX Key Location/Qualifiers
 FH Misc-difference 589
 FT /note= "Wild-type Thr substituted by Glu"
 XX US2002031826-A1.
 XX 14-MAR-2002.
 PD 14-MAR-2002.
 PF 19-DEC-2000; 2000US-0740274.
 XX 11-DEC-1998; 98US-0210361.
 PR 07-JUN-1995; 95US-0478704.
 PR 07-JUN-1995; 95US-0482711.
 PR 07-JUN-1995; 95US-0485243.
 PR 16-JAN-1998; 98US-0007999.
 PR 16-JAN-1998; 98US-0008172.
 PR 20-JAN-1998; 98US-0009620.
 XX (NICH/) NICHOLS S E.
 XX Nichols SE;
 PI WPI; 2002-414332/44.
 DR WPI; 2002-414332/44.
 XX Glucosyltransferase B or D protein useful for producing a glucan useful
 PT as substitutes for and additions to modified starch and latexes in
 PT paper manufacture, comprises mutations in specific positions -
 XX Claim 36; Page -; 44pp; English.

The invention an isolated protein comprising a glucosyltransferase
 (GTF) B polypeptide having changes at position from 1448Y, D457N,
 CC D567T, K1014T, D457N/D567T, D457N/D571K, D567T/D571K,
 CC D567T/D571K/K1014T, 1448Y/D457N/D567T/D571K/K719Q/K1014T,
 CC Y169A/Y170A/Y171A, and K779Q or a GTF D polypeptide having
 CC changes at positions from T589D, T589E, N471D, N471D/T589D, and
 CC N471D/T589E. Also included are a glucan produced by the GTF mutant,
 CC an isolated polynucleotide which encodes P1 or P2, or its complementary
 CC polynucleotide, a ribonucleic acid sequence encoding the GTF mutant,
 CC an expression cassette comprising the polynucleotide operably linked to a
 CC promoter, a vector comprising the expression cassette, host cell
 CC introduced with the vector, a transgenic plant comprising the
 CC vector, a seed or tuber from the transgenic plant, a paper sizing and/or
 CC coating composition comprising a glucan produced in a plant transformed
 CC with a gene encoding the mutant GTF, wild type or, starch, a latex,

thermoplastic molecule or their combinations or glucan and starch where
 CC the glucan is produced in the amyloplast and/or vacuole or a maize line
 CC deficient in starch biosynthesis, transformed with a gene encoding a
 CC glucosyltransferase B or D enzyme, wild-type or mutant and a paper
 CC comprising the glucan (paper sizing/coating agent). The vector is useful
 CC for producing a glucan in a plant. The method comprises transforming a
 CC plant cell with the vector, growing the plant cell under plant growing
 CC conditions to produce a regenerated plant and inducing expression of the
 CC polynucleotide for a time sufficient to produce the glucan in the
 CC regenerated plant, where the vector contains a transit sequence from
 CC ribulose biphosphate carboxylase small subunit, waxy, brittle-1 and
 CC chlorophyll AB binding protein to produce a transgenic plant, and glucan
 CC is produced in the amyloplast of potato or the vacuole of sugar beet.
 CC Glucans are useful as substitutes for and additions to modified starch
 CC and latexes in paper manufacture unlike prior art techniques, which
 CC require input materials that produce chemical effluents, paper
 CC manufacture utilizing the glucan produced by GTF, which utilizes
 CC biologically produced input materials, is more cost-effective and
 CC environmentally friendly. Moreover, glucans also exhibit thermoplastic
 CC properties and impart gloss to the paper during coating step.
 CC The present sequence represents a GTFD mutant of the invention.
 CC Note: The present sequence is not shown in the specification but
 CC was created by the indexer using the GTFD sequence appearing as AAU98029
 CC and the information in claim 36.

SQ Sequence 1430 AA;

Query Match 100.0%; Score 116; DB 23; Length 1430;
 Best Local Similarity 100.0%; Pred. No. 2.6e-09;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 AINHLSTLEAWSNDPQYKND 21
 |||||
 495 AINHLSTLEAWSNDPQYKND 515

RESULT 4
 AAU98043
 ID AAU98043 standard; Protein; 1430 AA.
 XX AAU98043;
 AC AAU98043;
 XX 27-AUG-2002 (first entry)
 DT 27-AUG-2002 (first entry)
 DE S. mutans glucosyltransferase GTFD mutant N471D.
 XX Glucosyltransferase; GTFD; transgenic plant; paper sizing;
 KM coating composition; glucan; starch; latex; thermoplastic molecule;
 KW amyloplast; vacuole; paper manufacture; mutant; mutain.
 XX Streptococcus mutans.
 OS Synthetic.
 XX Key Location/Qualifiers
 FH Misc-difference 471
 FT /note= "Wild-type Asn substituted by Asp"
 XX US2002031826-A1.
 XX 14-MAR-2002.
 PD 14-MAR-2002.
 PF 19-DEC-2000; 2000US-0740274.
 XX 11-DEC-1998; 98US-0210361.
 PR 07-JUN-1995; 95US-0478704.
 PR 07-JUN-1995; 95US-0482711.
 PR 07-JUN-1995; 95US-0485243.
 PR 16-JAN-1998; 98US-0007999.
 PR 16-JAN-1998; 98US-0008172.
 PR 20-JAN-1998; 98US-0009620.
 XX (NICH/) NICHOLS S E.
 XX Nichols SE;
 PI WPI; 2002-414332/44.
 DR WPI; 2002-414332/44.
 XX Glucosyltransferase B or D protein useful for producing a glucan useful
 PT as substitutes for and additions to modified starch and latexes in
 PT paper manufacture, comprises mutations in specific positions -
 XX Claim 36; Page -; 44pp; English.

XX Nichols SE;
XX WPI: 2002-414332/44.
XX
XX Glucosyltransferase B or D protein useful for producing a glucan useful
PT as substitutes for and additions to modified starch and latexes in
PT paper manufacture, comprises mutations in specific positions -
XX
XX Claim 36, Page -: 44pp; English.

XX The invention an isolated protein comprising a glucosyltransferase
CC (GTF) B polypeptide having changes at position from I448V, D457N,
CC D567T, K1014T, D457N/D567T, D457N/D571K, D567T/D571K,
CC D567T/D571K/K1014T, I448V/D457N/D567T/D571K/K779Q/K1014T,
CC Y169A/Y170A/Y171A, and K779Q or a GTF D polypeptide having
CC changes at positions from T589D, T589E, N471D, N471D/T589D, and
CC N471D/T589E. Also included are a glucan produced by the GTF mutant,
CC an isolated polynucleotide which encodes P1 or P2, or its complementary
CC polynucleotide, a ribonucleic acid sequence encoding the GTF mutant,
CC an expression cassette comprising the expression cassette, host cell
CC promoter, a vector comprising the expression cassette, host cell
CC introduced with the vector, a transgenic plant comprising the
CC vector, a seed or tuber from the transgenic plant, a paper string and/or
CC coating composition comprising a glucan produced in a plant transformed
CC with a gene encoding the mutant GTF, wild-type or, starch, a latex,
CC thermoplastic molecule or their combinations or glucan and starch where
CC the glucan is produced in the amyloplast and/or vacuole or a maize line
CC deficient in starch biosynthesis, transformed with a gene encoding a
CC glucosyltransferase B or D enzyme, wild-type or mutant and a paper
CC comprising the glucan (paper sizing/coating agent). The vector is useful
CC for producing a glucan in a plant. The method comprises transforming a
CC plant cell with the vector, growing the plant cell under plant growing
CC conditions to produce a regenerated plant and inducing expression of the
CC polynucleotide for a time sufficient to produce the glucan in the
CC regenerated plant, where the vector contains a transit sequence from
CC ribulose biphosphate carboxylase small subunit, waxy, brittle-1 and
CC chlorophyll AB binding protein to produce a transgenic plant, and glucan
CC is produced in the amyloplast of potato or the vacuole of sugar beet.
CC Glucans are useful as substitutes for and additions to modified starch
CC and latexes in paper manufacture. Unlike prior art techniques, which
CC require input materials that produce chemical effluents, paper
CC manufacture utilizing the glucan produced by GTF, which utilizes
CC biologically produced input materials, is more cost-effective and
CC environmentally friendly. Moreover, glucans also exhibit thermoplastic
CC properties and impart gloss to the paper during coating step.
CC The present sequence represents a GTF mutant of the invention.
CC Note: The present sequence is not shown in the specification but
CC was created by the indexer using the GTF sequence appearing as AAU98029
CC and the information in claim 36.
XX
XX

XX Sequence 1430 AA;
XX
XX Query Match 100.0%; Score 116; DB 23; Length 1430;
XX Best Local Similarity 100.0%; Pred. No. 2.6e-09;
XX Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AINHTSLERMSDNDPOYKND 21
DB 495 AINHTSLERMSDNDPOYKND 515

XX RESULT 5
XX AAU98044
XX AAU98044 standard; Protein: 1430 AA.
XX
XX AAU98044;
XX
XX 27-AAU9-2002 (first entry)
XX
XX S. mutans glucosyltransferase GTFD mutant N471D/T589D.
XX
XX Glucosyltransferase; GTFD; transgenic plant; paper sizing;
XX coating composition; glucan; starch; latex; thermoplastic molecule;
XX

KW amyloplast; vacuole; paper manufacture; mutant; mutain.
XX
XX Streptococcus mutans.
OS Synthetic.
OS
XX Key Location/Qualifiers
FH Misc-difference 471
FT Misc-difference 471
FT /note= "wild-type Asn substituted by Asp"
FT Misc-difference 589
FT /note= "wild-type Thr substituted by Asp"
XX
XX US2002031826-A1.
XX
XX 14-MAR-2002.
XX
XX 19-DEC-2000, 2000US-0740274.
XX
XX 11-DEC-1998, 98US-0210361.
XX 07-JUN-1995; 95US-0478704.
XX 07-JUN-1995; 95US-0482711.
XX 07-JUN-1995; 95US-0485243.
XX 16-JAN-1998; 98US-0007999.
XX 16-JAN-1998; 98US-0008172.
XX 20-JAN-1998; 98US-0009620.

XX (NICH/) NICHOLS S E.

XX Nichols SE;

XX WPI: 2002-414332/44.

XX Glucosyltransferase B or D protein useful for producing a glucan useful
PT as substitutes for and additions to modified starch and latexes in
PT paper manufacture, comprises mutations in specific positions -
XX
XX Claim 36, Page -: 44pp; English.

XX The invention an isolated protein comprising a glucosyltransferase
CC (GTF) B polypeptide having changes at position from I448V, D457N,
CC D567T, K1014T, D457N/D567T, D457N/D571K, D567T/D571K,
CC D567T/D571K/K1014T, I448V/D457N/D567T/D571K/K779Q/K1014T,
CC Y169A/Y170A/Y171A, and K779Q or a GTF D polypeptide having
CC changes at positions from T589D, T589E, N471D, N471D/T589D, and
CC N471D/T589E. Also included are a glucan produced by the GTF mutant,
CC an isolated polynucleotide which encodes P1 or P2, or its complementary
CC polynucleotide, a ribonucleic acid sequence encoding the GTF mutant,
CC an expression cassette comprising the expression cassette, host cell
CC promoter, a vector comprising the expression cassette, host cell
CC introduced with the vector, a transgenic plant comprising the
CC vector, a seed or tuber from the transgenic plant, a paper string and/or
CC coating composition comprising a glucan produced in a plant transformed
CC with a gene encoding the mutant GTF, wild-type or, starch, a latex,
CC thermoplastic molecule or their combinations or glucan and starch where
CC the glucan is produced in the amyloplast and/or vacuole or a maize line
CC deficient in starch biosynthesis, transformed with a gene encoding a
CC glucosyltransferase B or D enzyme, wild-type or mutant and a paper
CC comprising the glucan (paper sizing/coating agent). The vector is useful
CC for producing a glucan in a plant. The method comprises transforming a
CC plant cell with the vector, growing the plant cell under plant growing
CC conditions to produce a regenerated plant and inducing expression of the
CC polynucleotide for a time sufficient to produce the glucan in the
CC regenerated plant, where the vector contains a transit sequence from
CC ribulose biphosphate carboxylase small subunit, waxy, brittle-1 and
CC chlorophyll AB binding protein to produce a transgenic plant, and glucan
CC is produced in the amyloplast of potato or the vacuole of sugar beet.
CC Glucans are useful as substitutes for and additions to modified starch
CC and latexes in paper manufacture. Unlike prior art techniques, which
CC require input materials that produce chemical effluents, paper
CC manufacture utilizing the glucan produced by GTF, which utilizes
CC biologically produced input materials, is more cost-effective and
CC environmentally friendly. Moreover, glucans also exhibit thermoplastic
CC properties and impart gloss to the paper during coating step.
CC The present sequence represents a GTF mutant of the invention.

CC Note: The present sequence is not shown in the specification but
 CC was created by the indexer using the GTFD sequence appearing as AAU98029
 CC and the information in claim 36.

CC Sequence 1430 AA;

Query Match 100.0%; Score 116; DB 23; Length 1430;
 Best Local Similarity 100.0%; Pred. No. 2.6e-09;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AINHLSTLEAWSDNDPOYNKD 21
 |||||
 DB 495 AINHLSTLEAWSDNDPOYNKD 515

RESULT 6
 AAU98045
 ID AAU98045 standard; Protein; 1430 AA.

XX AAU98045;

DT 27-AUG-2002 (first entry)

XX S. mutans glucosyltransferase GTFD mutant N471D/T589E.

XX Glucosyltransferase; GTFD; transgenic plant; paper sizing;
 KM coating composition; glucan; starch; latex; thermoplastic molecule;
 XX amyloplast; vacuole; paper manufacture; mutant; mutain.

XX Streptococcus mutans.
 OS Synthetic.

XX Key Location/Qualifiers

FT Misc-difference 471 /note= "Wild-type Asn substituted by Asp"

FT Misc-difference 589 /note= "Wild-type Thr substituted by Glu"

XX US2002031826-A1.

XX 14-MAR-2002.

XX 19-DEC-2000; 2000US-0740274.

XX 11-DEC-1998; 98US-0210361.

XX 07-JUN-1995; 95US-0478704.

XX 07-JUN-1995; 95US-0462711.

XX 07-JUN-1995; 95US-0485243.

XX 16-JAN-1998; 98US-0007999.

XX 16-JAN-1998; 98US-0008172.

XX 20-JAN-1998; 98US-0009620.

XX (NICH/) NICHOLS S E.

XX Nichols SE;

XX WPI; 2002-414332/44.

XX Claim 36; Page -; 44pp; English.

XX The invention an isolated protein comprising a glucosyltransferase
 CC (GTF) B polypeptide having changes at position from 1448V, D457N,
 CC D567T, K1014T, D457N/D567T, D457N/D571K, D567T/D571K,
 CC D567T/D571K/K1014T, 1448V/D457N/D567T/D571K/K799Q/K1014T,
 CC Y169A/Y170A/Y171A, and K779Q or a GTF D polypeptide having
 CC changes at positions from T589D, T589E, N471D, N471D/T589D, and
 CC N471D/T589E. Also included are a glucan produced by the GTF mutant,
 CC an isolated polynucleotide which encodes P1 or P2, or its complementary
 CC polynucleotide, a ribonucleic acid sequence encoding the GTF mutant.

CC an expression cassette comprising the polynucleotide operably linked to a
 CC promoter, a vector comprising the expression cassette, host cell
 CC introduced with the vector, a transgenic plant comprising the
 CC vector, a seed or tuber from the transgenic plant, a paper sizing and/or
 CC coating composition comprising a glucan produced in a plant transformed
 CC with a gene encoding the mutant GTF, wild type or, starch, a latex,
 CC thermoplastic molecule or their combinations or glucan and starch where
 CC the glucan is produced in the amyloplast and/or vacuole or a maize line
 CC deficient in starch biosynthesis, transformed with a gene encoding a
 CC glucosyltransferase B or D enzyme, wild-type or mutant and a paper
 CC comprising the glucan (paper sizing/coating agent). The vector is useful
 CC for producing a glucan in a plant. The method comprises transforming a
 CC plant cell with the vector, growing the plant cell under plant growing
 CC conditions to produce a regenerated plant and inducing expression of the
 CC polynucleotide for a time sufficient to produce the glucan in the
 CC regenerated plant, where the vector contains a transit sequence from
 CC ribulose biphosphate carboxylase small subunit, waxy, brittle-1 and
 CC chlorophyll a binding protein to produce a transgenic plant, and glucan
 CC is produced in the amyloplast of potato or the vacuole of sugar beet.
 CC Glucans are useful as substitutes for and additions to modified starch
 CC and latexes in paper manufacture. Unlike prior art techniques, which
 CC require input materials that produce chemical effluents, paper
 CC manufacture utilizing the glucan produced by GTF, which utilizes
 CC biologically produced input materials, is more cost-effective and
 CC environmentally friendly. Moreover, glucans also exhibit thermoplastic
 CC properties and impart gloss to the paper during coating step.
 CC The present sequence represents a GTFD mutant of the invention.
 CC Note: The present sequence is not shown in the specification but
 CC was created by the indexer using the GTFD sequence appearing as AAU98029
 CC and the information in claim 36.

XX Sequence 1430 AA;

Query Match 100.0%; Score 116; DB 23; Length 1430;
 Best Local Similarity 100.0%; Pred. No. 2.6e-09;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AINHLSTLEAWSDNDPOYNKD 21
 |||||
 DB 495 AINHLSTLEAWSDNDPOYNKD 515

RESULT 7

AA91047 standard; Protein; 1577 AA.

XX AA91047;

DT 22-MAY-1996 (first entry)

XX Alpha-D-glucosyltransferase.

XX Alpha-D-glucosyltransferase; primer-independent; soluble glucan;

XX sucrose; transgenic plant; cloning; Escherichia coli;

XX phage lambda-C13; vector; plasmid pSG501; plasmid pSG502;

XX gene transfer; crop improvement; storage carbohydrate; pasture;

XX feedstuff; senescence; dextran; binder; food; pharmaceutical.

XX Streptococcus salivarius strain ATCC 25975.

XX W09606173-A1.

XX 29-FEB-1996.

XX 24-AUG-1995; 95WO-AU00527.

XX 24-AUG-1994; 94AU-0007643.

XX (GTFE/) GIFFARD P M.
 XX (JACQ/) JACQUES N A.
 XX (STMP/) SIMPSON C L.
 XX Giffard PM, Jacques NA, Simpson CL;

XX DR WPI, 1996-151376/15.
 XX DR N-PSDB; AAT13139.
 XX
 PT Plants contg. new bacterial DNA encoding glucosyl transferase
 PT activity - retain higher levels of stored carbohydrate(s) in a form
 PT readily digestible by ruminants
 XX
 PS Claim 4; Page 16-20; 31pp; English.
 XX
 CC The sequence represents an alpha-D-glucosyltransferase from
 CC Streptococcus salivarius. The enzyme is primer-independent, and
 CC produces soluble glucan from sucrose. A gene encoding the enzyme
 CC may be cloned and expressed in *Escherichia coli* using a subclone
 CC of phage lambda-C13, e.g. plasmid pSG501 or plasmid pSG502. The
 CC DNA may also be expressed in a transgenic plant, to improve the
 CC level of stored carbohydrate in a pasture plant which normally
 CC contains low levels, or to prevent degradation of stored carbohydrate
 CC during plant senescence. Dextran may be isolated from the plant, for
 CC use as a food binder or pharmaceutical additive. Primer independence
 CC ensures that the enzyme will be functional in plants. The glucan is
 CC poorly degraded in plants but easily degraded by bacteria in the rumen
 CC of grazing livestock.
 CC
 XX Sequence 1577 AA;
 SQ
 Query Match 77.6%; Score 90; DB 17; Length 1577;
 Best Local Similarity 85.7%; Pred. No. 4.1e-05;
 Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 OY 1 AINHLSTLEAWSNDPQYKXD 21
 Db 588 AIAHLSTLEAWSYNDHQYKXD 608
 RESULT 8
 AAU79285
 ID AAU79285 standard; Protein; 1017 AA.
 XX
 AC AAU79285;
 XX
 DT 13-AUG-2002 (first entry)
 XX
 DE Streptococcus mutans monoclonal antibody-related protein #2.
 XX
 KW Antibody; dental caries; water insoluble glucan synthetase;
 KW anti-caries; glucosyl transferase-B; immunotherapy.
 XX
 OS Streptococcus mutans.
 XX
 PN JP2002114709-A.
 PD
 XX 16-APR-2002.
 PF 04-OCT-2000; 2000JP-0304889.
 XX
 PR 04-OCT-2000; 2000JP-0304889.
 XX
 PA (UYNI-) UNIV NIPPON.
 XX
 DR WPI; 2002-448101/48.
 XX
 PT Anti-caries agent composed of a monoclonal antibody against an
 PT inhibitory enzyme against water insoluble glucan synthetase of glucosyl
 PT transferase-B (GTF-B) of Streptococcus mutans -
 XX
 PS Claim 4; Page 17-19; 28pp; Japanese.
 XX
 CC The invention relates to a monoclonal antibody against dental caries and
 CC an anti-caries agent composed of a monoclonal antibody produced by
 CC Streptococcus mutans, particularly mouse-hybridoma MHP126 (FERM P-17566)
 CC or mouse-hybridoma MHP136 (FERM P-17567), against an enzyme having
 CC inhibitive activity against water insoluble glucan synthetase of glucosyl

CC transferase-B. The monoclonal antibody specifically inhibits water
 CC insoluble glucan synthetase of Streptococcus mutans produced glucosyl
 CC transferase-B and is used in the immunotherapy of dental caries. This
 CC sequence represents a Streptococcus mutans monoclonal antibody-related
 CC protein.
 CC
 XX Sequence 1017 AA;
 SQ
 Query Match 64.7%; Score 75; DB 23; Length 1017;
 Best Local Similarity 71.4%; Pred. No. 0.0058;
 Matches 15; Conservative 1; Mismatches 5; Indels 0; Gaps 0;
 OY 1 AINHLSTLEAWSNDPQYKXD 21
 Db 447 ANDHLSTLEAWSNDPQYKXD 467
 RESULT 9
 AAU98027
 ID AAU98027 standard; Protein; 1475 AA.
 XX
 AC AAU98027;
 XX
 DT 27-AUG-2002 (first entry)
 XX
 DE S. mutans glucosyltransferase GTFB.
 XX
 KW Glucosyltransferase; GTFB; transgenic plant; paper sizing;
 KW coating composition; glucan; starch; latex; thermoplastic molecule;
 KW amyloplast; vacuole; paper manufacture.
 XX
 OS Streptococcus mutans.
 XX
 PN US2002031826-A1.
 PD
 XX 14-MAR-2002.
 PF 19-DEC-2000; 2000US-0740274.
 XX
 PR 11-DEC-1998; 98US-0210361.
 PR 07-JUN-1995; 95US-0478704.
 PR 07-JUN-1995; 95US-0482711.
 PR 07-JUN-1995; 95US-0485243.
 PR 16-JAN-1998; 98US-0007999.
 PR 16-JAN-1998; 98US-0008172.
 PR 20-JAN-1998; 98US-0009620.
 XX
 PA (NICH/) NICHOLS S E.
 XX
 PI Nichols SE;
 XX
 DR WPI; 2002-414332/44.
 DR N-PSDB; ABR52938.
 XX
 PT Glucosyltransferase B or D protein useful for producing a glucan useful
 PT as substitutes for and additions to modified starch and latexes in
 PT paper manufacture, comprises mutations in specific positions -
 XX
 PS Disclosure; Page 21-25; 44pp; English.
 XX
 CC The invention an isolated protein comprising a glucosyltransferase
 CC (GTF) B polypeptide having changes at position from I448V, D457N,
 CC D567T, K1014T, D457N/D567T, D457N/D571K, D567T/D571K,
 CC D567T/D571K/K1014T, I448V/D457N/D567T/D571K/K779Q/K1014T,
 CC Y169A/Y170A/Y171A, and K779Q or a GTF D polypeptide having
 CC changes at positions from T589D, T589S, N471D, N471D/T589D, and
 CC N471D/T589S. Also included are a glucan produced by the GTF mutant,
 CC an isolated polynucleotide which encodes P1 or P2, or its complementary
 CC polynucleotide, a ribonucleic acid sequence encoding the GTF mutant,
 CC an expression cassette comprising the polynucleotide operably linked to a
 CC promoter, a vector comprising the expression cassette, host cell
 CC introduced with the vector, a transgenic plant comprising the
 CC vector, a seed or tuber from the transgenic plant, a paper sizing and/or

coating composition comprising glucan produced in a plant transformed with a gene encoding the mutant GTF, wild type or, starch, a latex, thermoplastic molecule or their combinations or glucan and starch where the glucan is produced in the amyloplast and/or vacuole of a maize line deficient in starch biosynthesis, transformed with a gene encoding a glucosyltransferase B or D enzyme, wild-type or mutant and a paper comprising the glucan (paper sizing/coating agent). The vector is useful for producing a glucan in a plant. The method comprises transforming a plant cell with the vector, growing the plant cell under plant growing conditions to produce a regenerated plant and inducing expression of the polynucleotide for a time sufficient to produce the glucan in the regenerated plant, where the vector contains a transit sequence from ribulose biphosphate carboxylase small subunit, waxy, brittle-1 and chlorophyll AB binding protein to produce a transgenic plant, and glucose is produced in the amyloplast of potato or the vacuole of sugar beet. Glucans are useful as substitutes for and additions to modified starch and latexes in paper manufacture. Unlike prior art techniques, which require input materials that produce chemical effluents, paper manufacture utilizing the glucan produced by GTF, which utilizes biologically produced input materials, is more cost-effective and environmentally friendly. Moreover, glucans also exhibit thermoplastic properties and impart gloss to the paper during coating step. The present sequence represents GTFB.

SQ Sequence 1475 AA;

Query Match	64.7%	Score 75;	DB 23;	length 1475;
Best Local Similarity	71.4%	Pred. No. 0.0092;		
Matches 15;	Conservative	1;	Mismatches 5;	Indels 0;
			Gaps	0;

```

QY      1 AINHTSLLEAWMSDNDPQYNKD 21
      : ||||| |||||
DB      481 ANDHTSLLEAWMSDNDTPYLHD 501

```

RESULT 10
AAU98030
ID AAU98030 standard; Protein; 1475 AA

DT	27-AUG-2002	(first entry)
XX		
DE	S. mutans glucosyltransferase	GTFB mutant I448V

KM Glucosyltransferase; GTFB; transgenic plant; paper sizing;
KM coating composition; glucan; starch; latex; thermoplastic molecule
KM amyloplast; vacuole; paper manufacture; mutant; mutein.

OS Streptococcus mutans.
OS Synthetic.

EH	Key	Location/Qualifiers
FT	Misc-difference	448
FT		/note= "Wild-type Ile substituted by Val"

PN US2002031826-A1.

PD 14-MAR-2002

PF 19-DEC-2000; 2000US-0740274

PR 11-DEC-1998; 98US-0210361

PR 07-JUN-1995; 95US-0482/11

PR 16-JAN-1998; 98US-000/999
10 JAN 1998; 98US-0008173

PR 20-JAN-1998; 3005-0005820
XX

XX
FM
(INCL) / WYCHES D E.

7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844

PI Nichols SE,

XX WPI; 2002-414332/44.
DR

XX Glucosyltransferase B or D protein useful for producing a glucan useful
 XX as substrates for and additions to modified starch and latexes in
 XX paper manufacture, comprises mutations in specific positions -
 PS Claim 36; Page -; 44pp; English.

The invention an isolated protein comprising a glucosyltransferase (GTF) B polypeptide having changes at position from I448V, D45N, D567T, K1014T, D457N/D567T, D457N/D571K, D567T/D571K, D567T/D571K, I448V/D457N/D567T/D571K/K793Q/K1014T, Y169A/Y170A/Y171A, and K779Q or a GTF D polypeptide having changes at positions from T589D, T589E, N471D, N471D/T589D, and N471D/T589E. Also included are a glucan produced by the GTF mutant, an isolated polynucleotide which encodes P1 or P2, or its complementary polynucleotide, a ribonucleic acid sequence encoding the GTF mutant, an expression cassette comprising the polynucleotide operably linked to a promoter, a vector comprising the expression cassette, host cell introduced with the vector, a transgenic plant comprising the vector, a seed or tuber from the transgenic plant, a paper sizing and/or coating composition comprising a glucan produced in a plant transformed with a gene encoding the mutant GTF, wild type or, starch, a latex, thermoplastic molecule or their combinations and/or vacuole or a maize line the glucan is produced in the amyloplast and/or vacuole with a gene encoding a deficient in starch biosynthesis, transformed with a gene encoding a glucosyltransferase B or D enzyme, wild-type or mutant and a paper comprising the glucan (paper sizing/coating agent). The vector is useful for producing a glucan in a plant. The method comprises transforming a plant cell with the vector, growing the plant cell under plant growing conditions to produce a regenerated plant and inducing expression of the polynucleotide for a time sufficient to produce the glucan in the regenerated plant, where the vector contains a transit sequence from ribulose biphosphate carboxylase small subunit, waxy, brittle-1 and chlorophyll a/b binding protein to produce a transgenic plant, and glucan is produced in the amyloplast of potato or the vacuole of sugar beet. Glucans are useful as substitutes for and additions to modified starch and latexes in paper manufacture. Unlike prior art techniques, which require input materials that produce chemical effluents, paper manufacture utilizing the glucan produced by GTF, which utilizes biologically produced input materials, is more cost-effective and environmentally friendly. Moreover, glucans also exhibit thermoplastic properties and impart gloss to the paper during coating step. The present sequence represents a GTFB mutant of the invention. Note: The present sequence is not shown in the specification but was created by the indexer using the GTFB sequence appearing as AAU98027 and the information in claim 36.

Sequence	1475 AA
SQ	

Query Match	64.7%	Score 75;	DB 23;	Length 1475;
Best Local Similarity	71.4%;	Pred. No. 0.0092;		
Matches 15;	Conservative 1;	Mismatches 5;	Indels 0;	Gaps 0;

```

OY      1 AINHLSTLEAWSNDNDPPQYNKD 21
         | : ||||| ||||| |
Db      481 ANDHLSILEAWSNDNDTPYLLHD 501

```

RESULT 11
AAU98031
ID AAU98031 standard; Protein; 1475 AA.

AC AAU98031;

DT 27-AUG-2002 (first entry)

S. mutans glucosyltransferase GTFB mucaine protein.

glucosyltransferase; GIRD, craniogastro-pharyngeal diverticulum; composition: glucan; starch; latex; thermoplastic molecule;

any topographical

XX Streptococcus mutans.
OS Synthetic.
XX Key Location/Qualifiers
FH Misc-difference 457
FT /note= "Wild-type Asp substituted by Asn"
XX US2002031826-A1.
XX
XX
XX 14-MAR-2002.
XX
XX 19-DEC-2000; 2000US-0740274.
XX
XX 11-DEC-1998; 98US-0210361.
XX 07-JUN-1995; 95US-0478704.
XX 07-JUN-1995; 95US-0482711.
XX 07-JUN-1995; 95US-0485243.
XX 16-JAN-1998; 98US-0007999.
XX 16-JAN-1998; 98US-0008172.
XX 20-JAN-1998; 98US-0009620.
XX
XX (NICH/) NICHOLS S E.
XX
XX Nichols SE;
XX
XX MPI; 2002-414332/44.
XX
XX Glucosyltransferase B or D protein useful for producing a glucan useful
PT as substitutes for and additions to modified starch and latexes in
PT paper manufacture, comprises mutations in specific positions -
XX
XX Claim 36; Page -; 44pp; English.
XX
XX The invention an isolated protein comprising a glucosyltransferase
CC (GTF) B polypeptide having changes at position from 1448V, D457N,
CC D567T, K1014T, D457N/D567T, D457N/D571K, D567T/D571K,
CC D567T/D571K/K1014T, 1448V/D457N/D567T/D571K/K779Q/K1014T,
CC Y169A/Y170A/Y171A, and K779Q or a GTF D polypeptide having
CC changes at positions from T589D, T589E, N471D, N471D/T589D, and
CC N471D/T589E. Also included are a glucan produced by the GTF mutant,
CC an isolated polynucleotide which encodes F1 or F2, or its complementary
CC an isolated polynucleotide which encodes F1 or F2, or its complementary
CC polynucleotide, a ribonucleic acid sequence encoding the GTF mutant,
CC an expression cassette comprising the polynucleotide operably linked to a
CC promoter, a vector comprising the expression cassette, host cell
CC introduced with the vector, a transgenic plant comprising the
CC vector, a seed or tuber from the transgenic plant, a paper sizing and/or
CC coating composition comprising a glucan produced in a plant transformed
CC with a gene encoding the mutant GTF, wild type or, starch, a latex,
CC thermoplastic molecule or their combinations or glucan and starch where
CC the glucan is produced in the amyloplast and/or vacuole or a maize line
CC deficient in starch biosynthesis, transformed with a gene encoding a
CC glucosyltransferase B or D enzyme, wild-type or mutant and a paper
CC comprising the glucan (paper sizing/coating agent). The vector is useful
CC for producing a glucan in a plant. The method comprises transforming a
CC plant cell with the vector, growing the plant cell under paper-growing
CC conditions to produce a regenerated plant and inducing expression of the
CC polynucleotide for a time sufficient to produce the glucan in the
CC regenerated plant, where the vector contains a transit sequence from
CC ribulose biphosphate carboxylase small subunit, waxy, brittle-1 and
CC chlorophyll AB binding protein to produce a transgenic plant, and glucan
CC is produced in the amyloplast of potato or the vacuole of sugar beet.
CC Glucans are useful as substitutes for and additions to modified starch
CC and latexes in paper manufacture. Unlike prior art techniques, which
CC require input materials that produce chemical effluents, paper
CC manufacture utilizing the glucan produced by GTF, which utilizes
CC biologically produced input materials, is more cost-effective and
CC environmentally friendly. Moreover, glucans also exhibit thermoplastic
CC properties and impart gloss to the paper during coating step.
CC The present sequence represents a GTFB mutant of the invention.
CC Note: The present sequence is not shown in the specification but
CC was created by the indexer using the GTFB sequence appearing as AAU96027
CC and the information in claim 36.

XX SQ Sequence 1475 AA;
XX Query Match 64.7%; Score 75; DB 23; Length 1475;
XX Best Local Similarity 71.4%; Pred. No. 0.0092;
XX Matches 15; Conservative 1; Mismatches 5; Indels 0; Gaps 0;
XX
XX 1 AINHLSTLEAWSDNDPQYKND 21
XX :|||||
XX Db 481 ANDHLSTLEAWSDNDPPLHD 501
XX
XX
XX RESULT 12
XX ID AAU98032
XX AAU98032 standard; Protein; 1475 AA.
XX
XX AAU98032;
XX
XX 27-AUG-2002 (first entry)
XX
XX S. mutans glucosyltransferase GTFB mutant D567T.
XX
XX Glucosyltransferase; GTFB; transgenic plant; paper sizing;
XX coating composition; glucan; starch; latex; thermoplastic molecule;
XX amyloplast; vacuole; paper manufacture; mutant; mutain.
XX
XX Streptococcus mutans.
XX Synthetic.
XX
XX Key Location/Qualifiers
XX FH Misc-difference 567
XX FT /note= "Wild-type Asp substituted by Thr"
XX
XX US2002031826-A1.
XX
XX 14-MAR-2002.
XX
XX 19-DEC-2000; 2000US-0740274.
XX
XX 11-DEC-1998; 98US-0210361.
XX 07-JUN-1995; 95US-0478704.
XX 07-JUN-1995; 95US-0482711.
XX 07-JUN-1995; 95US-0485243.
XX 16-JAN-1998; 98US-0007999.
XX 16-JAN-1998; 98US-0008172.
XX 20-JAN-1998; 98US-0009620.
XX
XX (NICH/) NICHOLS S E.
XX
XX Nichols SE;
XX
XX MPI; 2002-414332/44.
XX
XX Glucosyltransferase B or D protein useful for producing a glucan useful
PT as substitutes for and additions to modified starch and latexes in
PT paper manufacture, comprises mutations in specific positions -
XX
XX Claim 36; Page -; 44pp; English.
XX
XX The invention an isolated protein comprising a glucosyltransferase
CC (GTF) B polypeptide having changes at position from 1448V, D457N,
CC D567T, K1014T, D457N/D567T, D457N/D571K, D567T/D571K,
CC D567T/D571K/K1014T, 1448V/D457N/D567T/D571K/K779Q/K1014T,
CC Y169A/Y170A/Y171A, and K779Q or a GTF D polypeptide having
CC changes at positions from T589D, T589E, N471D, N471D/T589D, and
CC N471D/T589E. Also included are a glucan produced by the GTF mutant,
CC an isolated polynucleotide which encodes F1 or F2, or its complementary
CC polynucleotide, a ribonucleic acid sequence encoding the GTF mutant,
CC an expression cassette comprising the polynucleotide operably linked to a
CC promoter, a vector comprising the expression cassette, host cell
CC introduced with the vector, a transgenic plant comprising the
CC vector, a seed or tuber from the transgenic plant, a paper sizing and/or
CC coating composition comprising a glucan produced in a plant transformed

CC with a gene encoding the mutant GTF, wild type or, starch, a latex,
 CC thermoplastic molecule or their combinations or glucan and starch where
 CC the glucan is produced in the amyloplast and/or vacuole or a maize line
 CC deficient in starch biosynthesis, transformed with a gene encoding a
 CC glucosyltransferase B or D enzyme, wild-type or mutant and a paper
 CC comprising the glucan (paper sizing/coating agent). The vector is useful
 CC for producing a glucan in a plant. The method comprises transforming a
 CC plant cell with the vector, growing the plant cell under plant growing
 CC conditions to produce a regenerated plant and inducing expression of the
 CC polynucleotide for a time sufficient to produce the glucan in the
 CC regenerated plant, where the vector contains a transit sequence from
 CC ribulose biphosphate carboxylase small subunit, waxy, brittle-1 and
 CC chlorophyll AB binding protein to produce a transgenic plant, and glucan
 CC is produced in the amyloplast of potato or the vacuole of sugar beet.
 CC Glucans are useful as substitutes for and additions to modified starch
 CC and latexes in paper manufacture. Unlike prior art techniques, which
 CC require input materials that produce chemical effluents, paper
 CC manufacture utilizing the glucan produced by GTF, which utilizes
 CC biologically produced input materials, is more cost-effective and
 CC environmentally friendly. Moreover, glucans also exhibit thermoplastic
 CC properties and impart gloss to the paper during coating step.
 CC The present sequence represents a GTFB mutant of the invention.
 CC Note: The present sequence is not shown in the specification but
 CC was created by the indexer using the GTFB sequence appearing as AAU98027
 CC and the information in claim 36.

XX Sequence 1475 AA;

Query Match 64.7%; Score 75; DB 23; Length 1475;

Best Local Similarity 71.4%; Pred. No. 0.0092; Indels 0; Gaps 0;

Matches 15; Conservative 1; Mismatches 5;

QY 1 AINHSILEAWSNDPQYKND 21

DB 481 ANDHSILEAWSNDPQYKND 501

RESULT 13
 AAU98033 standard; Protein, 1475 AA.

XX AAU98033;

DT 27-AUG-2002 (first entry)

XX S. mutans glucosyltransferase GTFB mutant K1014T.
 KW Glucosyltransferase; GTFB; transgenic plant; paper sizing;
 KW coating composition; glucan; starch; latex; thermoplastic molecule;
 KW amyloplast; vacuole; paper manufacture; mutant; mutain.

XX Streptococcus mutans.
 OS Synthetic.

XX Key Location/Qualifiers
 FH Misc-difference 1014

FT Note="Wild-type Lys substituted by Thr"

XX US2002031826-A1.

XX 14-MAR-2002.

XX 19-DEC-2000; 2000US-0740274.

XX 11-DEC-1998; 98US-0210361.

XX 07-JUN-1995; 95US-0478704.

XX 07-JUN-1995; 95US-0482711.

XX 07-JUN-1995; 95US-0485243.

XX 16-JAN-1998; 98US-0008172.

XX 20-JAN-1998; 98US-0009620.

XX (NICH/) NICHOLS S E.

XX Nichols SE;

DR WPI; 2002-414332/44.

XX Glucosyltransferase B or D protein useful for producing a glucan useful
 PT as substitutes for and additions to modified starch and latexes in
 PT paper manufacture, comprises mutations in specific positions -

XX Claim 36; Page -; 44pp; English.

XX The invention an isolated protein comprising a glucosyltransferase
 CC (GTF) B polypeptide having changes at position from 1448V, D457N,
 CC D567T, K1014T, D457N/D567T, D457N/D571K, D567T/D571K,
 CC D567T/D571K/K1014T, 1448V/D457N/D567T/D571K/K779Q/K1014T,
 CC Y169A/Y170A/Y171A, and K779Q or a GTF D polypeptide having
 CC changes at positions from T589D, T589E, N471D, N471D/T589D, and
 CC N471D/T589E. Also included are a glucan produced by the GTF mutant,
 CC an isolated polynucleotide which encodes p1 or p2, or its complementary
 CC polynucleotide, a ribonucleic acid sequence encoding the GTF mutant,
 CC an expression cassette comprising the expression cassette, host cell
 CC promoter, a vector comprising the expression cassette comprising the
 CC introduced with the vector, a transgenic plant comprising the
 CC vector, a seed or tuber from the transgenic plant, a paper sizing and/or
 CC coating composition comprising a glucan produced in a plant transformed
 CC with a gene encoding the mutant GTF, wild type or, starch, a latex,
 CC thermoplastic molecule or their combinations and/or vacuole or a maize line
 CC deficient in starch biosynthesis, transformed with a gene encoding a
 CC glucan is produced in the amyloplast and/or vacuole or a maize line
 CC comprising the glucan (paper sizing/coating agent). The vector is useful
 CC for producing a glucan in a plant. The method comprises transforming a
 CC plant cell with the vector, growing the plant cell under plant growing
 CC conditions to produce a regenerated plant and inducing expression of the
 CC polynucleotide for a time sufficient to produce the glucan in the
 CC regenerated plant, where the vector contains a transit sequence from
 CC ribulose biphosphate carboxylase small subunit, waxy, brittle-1 and
 CC chlorophyll AB binding protein to produce a transgenic plant, and glucan
 CC is produced in the amyloplast of potato or the vacuole of sugar beet.
 CC Glucans are useful as substitutes for and additions to modified starch
 CC and latexes in paper manufacture. Unlike prior art techniques, which
 CC require input materials that produce chemical effluents, paper
 CC manufacture utilizing the glucan produced by GTF, which utilizes
 CC biologically produced input materials, is more cost-effective and
 CC environmentally friendly. Moreover, glucans also exhibit thermoplastic
 CC properties and impart gloss to the paper during coating step.
 CC The present sequence represents a GTFB mutant of the invention.
 CC Note: The present sequence is not shown in the specification but
 CC was created by the indexer using the GTFB sequence appearing as AAU98027
 CC and the information in claim 36.

XX Sequence 1475 AA;

Query Match 64.7%; Score 75; DB 23; Length 1475;

Best Local Similarity 71.4%; Pred. No. 0.0092; Indels 0; Gaps 0;

Matches 15; Conservative 1; Mismatches 5;

QY 1 AINHSILEAWSNDPQYKND 21

DB 481 ANDHSILEAWSNDPQYKND 501

RESULT 14

AAU98034 standard; Protein, 1475 AA.

XX AAU98034;

DT 27-AUG-2002 (first entry)

XX S. mutans glucosyltransferase GTFB mutant D457N/D567T.

XX Glucosyltransferase; GTFB; transgenic plant; paper sizing;

coating composition; glucan; starch; latex; thermoplastic molecule;
 amyloplast; vacuole; paper manufacture; mutant; muten.
 Streptococcus mutans.
 Synthetic.
 Key Location/Qualifiers
 Misc-difference 457 /note= "Wild-type Asp substituted by Asn"
 Misc-difference 567 /note= "Wild-type Asp substituted by Thr"
 US2002031826-A1.
 14-MAR-2002.
 19-DEC-2000; 2000US-0740274.
 11-DEC-1998; 98US-0210361.
 07-JUN-1995; 95US-0478704.
 07-JUN-1995; 95US-0482711.
 07-JUN-1995; 95US-0485243.
 16-JAN-1998; 98US-0007999.
 16-JAN-1998; 98US-0008172.
 20-JAN-1998; 98US-0009620.
 (NICH/) NICHOLS S E.
 Nichols SE;
 WPI; 2002-414332/44.
 Glucosyltransferase B or D protein useful for producing a glucan useful
 as substitutes for and additions to modified starch and latexes in
 paper manufacture, comprises mutations in specific positions -
 Claim 36; Page -; 44pp; English.
 The invention an isolated protein comprising a glucosyltransferase
 (GTF) B polypeptide having changes at position from 1448V, D457N,
 D567T, K1014T, D457N/D567T, D457N/D571K, D567T/D571K,
 D567T/D571K/K1014T, 1448V/D457N/D571K/K779Q/K1014T,
 Y169A/Y170A/Y171A, and K779Q or a GTF D polypeptide having
 changes at positions from T589D, T589E, N471D, N471D/T589D, and
 N471D/T589E. Also included are a glucan produced by the GTF mutant,
 an isolated polynucleotide which encodes P1 or P2, or its complementary
 polynucleotide, a ribonucleic acid sequence encoding the GTF mutant,
 an expression cassette comprising the polynucleotide operably linked to a
 promoter, a vector comprising the expression cassette, host cell
 introduced with the vector, a transgenic plant comprising the
 vector, a seed or tuber from the transgenic plant, a paper sizing and/or
 coating composition comprising a glucan produced in a plant transformed
 with a gene encoding the mutant GTF, wild type or, starch, a latex,
 thermoplastic molecule or their combinations or glucan and starch where
 the glucan is produced in the amyloplast and/or vacuole or a maize like
 deficient in starch biosynthesis, transformed with a gene encoding a
 glucosyltransferase B or D enzyme, wild-type or mutant and a paper
 comprising the glucan (paper sizing/coating agent). The vector is useful
 for producing a glucan in a plant. The method comprises transforming a
 plant cell with the vector, growing the plant cell under plant growing
 conditions to produce a regenerated plant and inducing expression of the
 polynucleotide for a time sufficient to produce the glucan in the
 regenerated plant, where the vector contains a transit sequence from
 ribulose biphosphate carboxylase small subunit, waxy, brittle-1 and
 chlorophyll AB binding protein to produce a transgenic plant, and glucan
 is produced in the amyloplast of potato or the vacuole of sugar beet.
 Glucans are useful as substitutes for and additions to modified starch
 and latexes in paper manufacture. Unlike prior art techniques, which
 require input materials that produce chemical effluents, paper
 manufacture utilizing the glucan produced by GTF, which utilizes
 biologically produced input materials, is more cost-effective and
 environmentally friendly. Moreover, glucans also exhibit thermoplastic
 properties and impart gloss to the paper during coating step.

The present sequence represents a GTFB mutant of the invention.
 Note: The present sequence is not shown in the specification but
 was created by the indexer using the GTFB sequence appearing as AAU98027
 and the information in claim 36.
 SO Sequence 1475 AA;
 Query Match 64.7%; Score 75; DB 23; Length 1475;
 Best Local Similarity 71.4%; Pred. No. 0.0092;
 Matches 15; Conservative 1; Mismatches 5; Indels 0; Gaps 0;
 QY 1 AINHLISLEAMSNDPQYNKD 21
 Db 481 ANDHLISLEAMSNDPQLHD 501
 RESULT 15
 AAU98035
 ID AAU98035 standard; Protein; 1475 AA.
 AC AAU98035;
 DT 27-AUG-2002 (first entry)
 XX S. mutans glucosyltransferase GTFB mutant D457N/D571K.
 XX Glucosyltransferase; GTFB; transgenic plant; paper sizing;
 KM coating composition; glucan; starch; latex; thermoplastic molecule;
 KM amyloplast; vacuole; paper manufacture; mutant; muten.
 XX Streptococcus mutans.
 OS Synthetic.
 XX Key Location/Qualifiers
 FH Misc-difference 457 /note= "Wild-type Asp substituted by Asn"
 FT Misc-difference 571 /note= "Wild-type Asp substituted by Lys"
 FT Misc-difference 571 /note= "Wild-type Asp substituted by Lys"
 XX US2002031826-A1.
 XX 14-MAR-2002.
 XX 19-DEC-2000; 2000US-0740274.
 XX 11-DEC-1998; 98US-0210361.
 PR 07-JUN-1995; 95US-0478704.
 PR 07-JUN-1995; 95US-0482711.
 PR 07-JUN-1995; 95US-0485243.
 PR 16-JAN-1998; 98US-0007999.
 PR 16-JAN-1998; 98US-0008172.
 PR 20-JAN-1998; 98US-0009620.
 PA (NICH/) NICHOLS S E.
 PI Nichols SE;
 WPI; 2002-414332/44.
 Glucosyltransferase B or D protein useful for producing a glucan useful
 as substitutes for and additions to modified starch and latexes in
 paper manufacture, comprises mutations in specific positions -
 Claim 36; Page -; 44pp; English.
 The invention an isolated protein comprising a glucosyltransferase
 (GTF) B polypeptide having changes at position from 1448V, D457N,
 D567T, K1014T, D457N/D567T, D457N/D571K, D567T/D571K,
 D567T/D571K/K1014T, 1448V/D457N/D571K/K779Q/K1014T,
 Y169A/Y170A/Y171A, and K779Q or a GTF D polypeptide having
 changes at positions from T589D, T589E, N471D, N471D/T589D, and
 N471D/T589E. Also included are a glucan produced by the GTF mutant,
 an isolated polynucleotide which encodes P1 or P2, or its complementary

CC polynucleotide, a ribonucleic acid sequence encoding the GFP mutant,
CC an expression cassette comprising the polynucleotide operably linked to a
CC promoter, a vector comprising the expression cassette, most cell
CC introduced with the vector, a transgenic plant comprising the
CC vector, a seed or tuber from the transgenic plant, a paper sizing and/or
CC coating composition comprising a glucan produced in a plant transformed
CC with a gene encoding the mutant GFP, wild type or, starch, a latex,
CC thermoplastic molecule or their combinations or glucan and starch where
CC the glucan is produced in the amyloplast and/or vacuole or a maize line
CC deficient in starch biosynthesists, transformed with a gene encoding a
CC glucosyltransferase B or D enzyme, wild-type or mutant and a paper
CC comprising the glucan (paper sizing/coating agent). The vector is useful
CC for producing a glucan in a plant. The method comprises transforming a
CC plant cell with the vector, growing the plant cell under plant growing
CC conditions to produce a regenerated plant and inducing expression of the
CC polynucleotide for a time sufficient to produce the glucan in the
CC regenerated plant, where the vector contains a transit sequence from
CC ribulose biphosphate carboxylase small subunit, waxy, brittle-1 and
CC chlorophyll AB binding protein to produce a transgenic plant, and glucan
CC is produced in the amyloplast of potato or the vacuole of sugar beet.
CC Glucans are useful as substitutes for and additions to modified starch
CC and latexes in paper manufacture. Unlike prior art techniques, which
CC require input materials that produce chemical effluents, paper
CC manufacture utilizing the glucan produced by GFP, which utilizes
CC biologically produced input materials, is more cost-effective and
CC environmentally friendly. Moreover, glucans also exhibit thermoplastic
CC properties and impart gloss to the paper during coating step.
CC The present sequence represents a GFP mutant of the invention.
CC Note: The present sequence is not shown in the specification but
CC was created by the indexer using the GFPB sequence appearing as ABU98027
CC and the information in claim 36.

SQ Sequence 1475 AA;

Query Match	64.7%;	Score 75;	DB 23;	Length 1475;
Best Local Similarity	71.4%;	Pred. No.	0.0092;	

Matches 15; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 AINHLSLEAWSNDNDPQYNKD 21

Db 481 ANDHLSLEAWSNDNDTPYLHD 501

Search completed: November 13, 2003, 09:38:26
Job time : 38.5166 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 13, 2003, 09:31:40 ; Search time 13.0379 Seconds
(without alignments)
154.898 Million cell updates/sec

Title: US-09-290-049a-11

Perfect score: 116

Sequence: 1 AINHSLILEAMSDNDPQYKND 21

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	116	100.0	1431	2 A45866	dextranucrase (EC
2	90	77.6	1577	2 T30858	glucosyltransferas
3	85	73.3	1449	2 T30857	glucosyltransferas
4	85	73.3	1449	2 T30852	glucosyltransferas
5	81	69.8	1365	2 A41483	glucosyltransferas
6	78	67.2	1508	2 T31098	probable dextranu
7	75	64.7	1475	2 B33135	gtfB protein precu
8	74	63.8	1592	2 A38175	glucosyltransferas
9	68	58.6	1518	2 A44811	glucosyltransferas
10	67	57.8	1599	2 S22737	glucosyltransferas
11	66	56.9	1375	2 JT0345	dextranucrase (EC
12	51	44.0	429	2 AH0630	4-hydroxyphenylac
13	49	42.2	632	2 T46504	hypothetical prote
14	47.5	40.9	369	2 AE2493	hypothetical prote
15	47	40.5	175	2 C86205	hypothetical prote
16	47	40.5	403	2 B70961	probable esterase
17	47	40.5	418	1 FOXRL2	sigma 2 protein -
18	47	40.5	445	2 H75360	cyclochrome P450 -
19	47	40.5	491	2 A86824	sensor protein kin
20	46	39.7	418	1 FOXRL5	sigma 2 protein -
21	46	39.7	420	2 T05877	hypothetical prote
22	46	39.7	623	2 T35377	probable membrane
23	46	39.7	706	2 D70178	PTS system, fructo
24	46	39.7	703	2 A48764	calpain (EC 3.4.22
25	45.5	39.2	250	2 G72495	probable polysulf
26	45	38.8	194	2 H72037	conserved hypothet
27	45	38.8	194	2 C86586	CT67 hypothetical
28	45	38.8	249	2 F95957	probable transcrip
29	45	38.8	344	2 T21604	hypothetical prote

30	45	38.8	440	2 C84265	adenylosuccinate s
31	45	38.8	504	2 S51942	prunin 2 precursor
32	45	38.8	637	2 T00548	hypothetical prote
33	45	38.8	644	2 A97268	methionyl-tRNA syn
34	45	38.8	657	2 S25184	cspl protein - Cor
35	45	38.8	1475	2 T29809	hypothetical prote
36	45	38.8	2358	2 T39569	probable alpha-glu
37	45	38.8	2371	2 T43432	alpha-glucan synth
38	44.5	38.4	188	2 A75382	hypothetical prote
39	44.5	38.4	236	2 S54428	fimbrial protein h
40	44.5	38.4	241	1 S24978	fimbrial protein h
41	44	37.9	122	2 AD1009	conserved hypothet
42	44	37.9	291	2 AB3235	nitrotriacetate
43	44	37.9	312	2 S67667	hypothetical prote
44	44	37.9	335	2 C85642	hypothetical prote
45	44	37.9	335	2 F90781	hypothetical prote

ALIGNMENTS

RESULT 1

A45866 dextranucrase (EC 2.4.1.5) precursor - Streptococcus mutans

C/Species: Streptococcus mutans

C/Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 07-May-1999

C/Accession: A45866

R/Honda, O.; Kato, C.; Kuramitsu, H.K.

J. Gen. Microbiol. 136, 2099-2105, 1990

A/Title: Nucleotide sequence of the Streptococcus mutans gtfD gene encoding the glucosyl

A/Accession: A45866

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-1431 <HON>

A/Cross-references: GB:M29296

C/Superfamily: cpl repeat homology

C/Keywords: glucosyltransferase; hexosyltransferase

F/181-201/Domain: cpl repeat homology <CP1>

F/1127-1146/Domain: cpl repeat homology <CP2>

F/1192-1211/Domain: cpl repeat homology <CP3>

F/1257-1276/Domain: cpl repeat homology <CP4>

F/1277-1297/Domain: cpl repeat homology <CP5>

F/1321-1340/Domain: cpl repeat homology <CP6>

F/1341-1361/Domain: cpl repeat homology <CP7>

F/1385-1404/Domain: cpl repeat homology <CP7>

Query Match Best Local Similarity 100.0%; Score 116; DB 2; Length 1431;

Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AINHSLILEAMSDNDPQYKND 21
DB 495 AINHSLILEAMSDNDPQYKND 515

RESULT 2

T30858 glucosyltransferase - Streptococcus salivarius

C/Species: Streptococcus salivarius

C/Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 22-Oct-1999

C/Accession: T30858

R/Simpson, C.L.; Giffard, P.M.; Jacques, N.A.

Infect. Immun. 63, 609-621, 1995

A/Title: Streptococcus salivarius ATCC 25975 possesses at least two genes coding for prim

A/Reference number: Z20909; MUID:95122197; PMID:7822030

A/Accession: T30858

A/Status: preliminary; translated from GB/EMBL/DDBJ

A/Molecule type: DNA

A/Residues: 1-1577 <SIM>

A/Cross-references: EMBL:L35928; NID:G662380; PID:G662381; PIDN:AA041413.1

C/Genetics:

A/Gene: gtfm

```

Query Match 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
Best Local Similarity 85.7%; Pred. No. 2,4e-05;
Matches 16; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

OY      1 AINHLSTLEAWSNDNDPOYNKD 21
          ||| ||||| ||||| |||||
Db       588 AIAHLSILEAWSNHDQYNKD 608

RESULT 3
glucosyltransferase - Streptococcus salivarius
C:Species: Streptococcus salivarius
C>Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 22-Oct-1995
C:Accession: J30857
R:Simpson, C.T.; Giffard, P.M.; Jacques, N.A.
Infect. Immun. 63, 609-621, 1995
A>Title: Streptococcus salivarius ATCC 25975 possesses at least two genes coding for pili
A:Reference number: Z20909; PMID:95122197; EMBL:7822030
A:Accession: J30857
A>Status: preliminary; translated from GB/EMBL/DDBT
A:Molecule type: DNA
A:Residues: 1-1449 <SIM>
A:Cross-references: EMBL:L35495; NID:g662378; PID:g662379; PIDN:AAC41412.1
C:Genetics:
A:Gene: gtfI

Query Match 16; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
Best Local Similarity 76.2%; Pred. No. 0.00013;
Matches 16; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

OY      1 AINHLSTLEAWSNDNDPOYNKD 21
          ||| ||||| ||||| |||||
Db       536 AIKHLSTLEAWSNHDAYNED 556

RESULT 4
glucosyltransferase N - Streptococcus salivarius (fragment)
C:Species: Streptococcus salivarius
C>Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 22-Oct-1999
C:Accession: J30552
R:Jaffe, R.I.
submitted to the EMBL Data Library, February 1998
A>Description: Streptococcus salivarius V1477 gtfN.
A:Reference number: Z20854
A:Accession: J30552
A>Status: preliminary; translated from GB/EMBL/DDBT
A:Molecule type: DNA
A:Residues: 1-1449 <UAF>
A:Cross-references: EMBL:AF049609; NID:g2935545; PID:g2935546; PIDN:AAC05156.1
C:Genetics:
A:Gene: gtfN

Query Match 16; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
Best Local Similarity 76.2%; Pred. No. 0.00013;
Matches 16; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

OY      1 AINHLSTLEAWSNDNDPOYNKD 21
          ||| ||||| ||||| |||||
Db       536 AIKHLSTLEAWSNHDAYNED 556

RESULT 5
glucosyltransferase (EC 2.4.1.-) gtfS precursor - Streptococcus sobrinus
C:Species: Streptococcus sobrinus
C>Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 15-Oct-1999
C:Accession: A41483
R:Gilmore, K.S.; Russell, R.R.B.; Ferretti, J.J.
Infect. Immun. 58, 2452-2458, 1990
A>Title: Analysis of the Streptococcus downii gtfS gene, which specifies a glucosyltransf
```

```

A:Reference number: A41483; MWID:90316665; PMID:2142479
A:Accession: A41483
A:Molecule type: DNA
A:Residues: 1-1365 <GIL>
A:Cross-references: GB:M30943; NID:g153652; PIDN:AAA6898.1; PID:g153653
C:Genetics:
A:Gene: gtfS
C:Superfamily: cpl repeat homology
C:Keywords: glycosyltransferase; hexosyltransferase

Query Match      69.8%; Score 81; DB 2; Length 1365;
Best Local Similarity 76.2%; Pred. No. 0.0005;
Matches 16; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Cy      1 AINHLSTLEAWSDNDPQYKND 21
      |||:|||||||:|||||
Db      467 AIDHLSTLEAWSGNDNDYKND 487

RESULT 6
T31098
Probable dextranucrase (EC 2.4.1.5), extracellular - Leuconostoc mesenteroides
C:Species: Leuconostoc mesenteroides
C:Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 11-May-2000
A:Accession: T31098
A:Monchois, V.; Remaud-Simeon, M.; Monsan, P.; Willemot, R.M.
FEMS Microbiol. Lett. 159, 307-315, 1998
A:Title: Cloning and sequencing of a gene coding for an extracellular dextranucrase
A:Reference number: Z20381; MWID:98164374; PMID:9503626
A:Accession: T31098
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1508 <MON>
A:Cross-references: EMBL:AF030129; NID:g2766611; PID:g2766612; PIDN:AAB5453.1
A:Experimental source: strain NRRL B-1299
C:Genetics:
A:Gene: dsxB
C:Function:
A:Description: produces dextran composed only of alpha(1-6) glucosidic bonds
C:Keywords: glycosyltransferase; hexosyltransferase

Query Match      67.2%; Score 78; DB 2; Length 1508;
Best Local Similarity 71.4%; Pred. No. 0.0016;
Matches 15; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

Cy      1 AINHLSTLEAWSDNDPQYKND 21
      |||:|||||||:|||||
Db      563 ANQHLSTLEAWSDNDPEYKND 583

RESULT 7
B33135
gtfB protein precursor - Streptococcus mutans
C:Species: Streptococcus mutans
C:Date: 23-Oct-1990 #sequence_revision 23-Oct-1990 #text_change 15-Oct-1999
A:Accession: B33135; A33128
A:Shiroza, T.; Ueda, S.; Kuramitsu, H.K.
J. Bacteriol. 169, 4263-4270, 1987
A:Title: Sequence analysis of the gtfB gene from Streptococcus mutans.
A:Reference number: A33135; MWID:87308013; PMID:3040685
A:Accession: B33135
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1475 <SHI>
A:Cross-references: GB:M17361; NID:g153639; PIDN:AAA89588.1; PID:g153640
R:Shiroza, T.; Ueda, S.; Kuramitsu, H.K.
submitted to the Protein Sequence Database, September 1990
A:Reference number: A33128
A:Accession: A33128
A:Status: preliminary; not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 1-171,173-641,'N',643-1475 <SH2>
A:Experimental source: strain GS-5

```

C:Superfamily: cpl repeat homology
F:1096-1115/Domain: cpl repeat homology <CP1>
F:1224-1243/Domain: cpl repeat homology <CP2>
F:1289-1308/Domain: cpl repeat homology <CP3>
F:1354-1373/Domain: cpl repeat homology <CP4>
F:1419-1438/Domain: cpl repeat homology <CP5>

Query Match 64.7%; Score 75; DB 2; Length 1475;
Best Local Similarity 71.4%; Pred. No. 0.0045;
Matches 15; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 AINHLSTLEAWSNDPQYKD 21
DB 481 AINHLSTLEAWSNDPQYKD 501

RESULT 8

glucosyltransferase precursor - Streptococcus sobrinus

C:Species: Streptococcus sobrinus
C:Date: 28-Aug-1992 #sequence_revision 28-Aug-1992 #text_change 15-Oct-1995

C:Accession: A38175
R:Abdo, H.; Msumura, T.; Kodama, T.; Ohta, H.; Fukui, K.; Kato, K.; Kagawa, H.

J. Bacteriol. 173, 989-996, 1991
A:Title: Peptide sequences for sucrose splitting and glucan binding within Streptococcus

A:Reference number: A38175; MUID:91123227; PMID:1704006

A:Accession: A38175
A:Status: Preliminary

A:Molecule type: DNA
A:Residues: 1-1592 <ABO>

A:Cross-references: GB:D90213; NID:g217032; PIDN:BA14241.1; PID:d1014946; PID:g217033

C:Superfamily: cpl repeat homology
F:1093-1112/Domain: cpl repeat homology <CP1>

F:1222-1241/Domain: cpl repeat homology <CP2>
F:1287-1306/Domain: cpl repeat homology <CP3>

F:1330-1351/Domain: cpl repeat homology <CP4>
F:1352-1371/Domain: cpl repeat homology <CP5>

F:1402-1420/Domain: cpl repeat homology <CP6>
F:1465-1484/Domain: cpl repeat homology <CP7>

F:1513-1532/Domain: cpl repeat homology <CP8>

Query Match 63.8%; Score 74; DB 2; Length 1592;
Best Local Similarity 66.7%; Pred. No. 0.007;
Matches 14; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 1 AINHLSTLEAWSNDPQYKD 21
DB 477 AINHLSTLEAWSNDPQYKD 497

glucosyltransferase (EC 2.4.1.-) I - Streptococcus salivarius

C:Species: Streptococcus salivarius
C:Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 15-Oct-1999

C:Accession: A44811; S22726; S28809
R:Giffard, P.M.; Simpson, C.L.; Milward, C.P.; Jacques, N.A.

J. Gen. Microbiol. 137, 2577-2593, 1991
A:Title: Molecular characterization of a cluster of at least two glucosyltransferase gen

A:Reference number: A44811; MUID:92148377; PMID:1838391

A:Accession: A44811
A:Molecule type: DNA

A:Residues: 1-1518 <GIF>
A:Cross-references: EMBL:211873; NID:g47526; PIDN:CAA77900.1; PID:g47527

C:Genetics: A:Note: sequence extracted from NCBI backbone (NCBI:81050, NCBI:81052)

A:Gene: gtfJ
C:Superfamily: cpl repeat homology
C:Keywords: glycosyltransferase; hexosyltransferase

F:1307-1326/Domain: cpl repeat homology <CP4>
Query Match 58.6%; Score 68; DB 2; Length 1518;
Best Local Similarity 63.2%; Pred. No. 0.056;

Matches 12; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 AINHLSTLEAWSNDPQYKD 19
DB 501 AINHLSTLEAWSNDPQYKD 519

RESULT 10

glucosyltransferase (EC 2.4.1.-) S - Streptococcus salivarius

C:Species: Streptococcus salivarius
C:Date: 16-Sep-1992 #sequence_revision 16-Sep-1992 #text_change 21-Jan-2000

C:Accession: S22737; S28810; E44811; S22727
R:Jacques, N.

submitted to the EMBL Data Library, March 1992
A:Reference number: S22737

A:Accession: S22737
A:Molecule type: DNA

A:Residues: 1-1599 <JAC>
A:Cross-references: EMBL:211873; NID:g47530; PIDN:CAA77898.1; PID:g47531

A:Experimental source: ATCC 25975
R:Giffard, P.M.; Simpson, C.L.; Milward, C.P.; Jacques, N.A.

J. Gen. Microbiol. 137, 2577-2593, 1991
A:Title: Molecular characterization of a cluster of at least two glucosyltransferase gen

A:Reference number: A44811; MUID:92148377; PMID:1838391

A:Accession: S28810
A:Molecule type: DNA

A:Residues: 1-51 <GIF>
A:Cross-references: EMBL:211873

C:Genetics: A:Gene: gtfK

C:Superfamily: cpl repeat homology
C:Keywords: glycosyltransferase; hexosyltransferase

F:1456-1475/Domain: cpl repeat homology <CP8>

Query Match 57.8%; Score 67; DB 2; Length 1599;
Best Local Similarity 66.7%; Pred. No. 0.084;
Matches 12; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 AINHLSTLEAWSNDPQY 18
DB 491 AINHLSTLEAWSNDPQY 508

RESULT 11

glucosyltransferase (EC 2.4.1.5) precursor - Streptococcus mutans (strain GS-5)

C:Species: Streptococcus mutans
C:Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 24-Sep-1999

C:Accession: J70345; C33135
R:Ueda, S.; Shiroza, T.; Kuramitsu, H.K.

Gene 69, 101-109, 1988
A:Title: Sequence analysis of the gtfC gene from Streptococcus mutans GS-5.

A:Reference number: J70345; MUID:89137980; PMID:2976010

A:Accession: J70345
A:Molecule type: DNA

A:Residues: 1-1375 <UDB>
A:Experimental source: GS-5

J. Bacteriol. 169, 4263-4270, 1987
A:Title: Sequence analysis of the gtfC gene from Streptococcus mutans.

A:Reference number: A33135; MUID:87508013; PMID:3040685

A:Accession: C33135
A:Molecule type: DNA

A:Residues: 1-349 <SHI>
A:Cross-references: GB:M17361

C:Genetics: A:Gene: gtfC

C:Function: A:Description: catalyzes the synthesis of both water-soluble and water-insoluble glucans
C:Superfamily: cpl repeat homology

C:Keywords: duplication; glycosyltransferase; hexosyltransferase
 F:1-34/Domain: signal sequence #status predicted <STO>
 F:35-1375/Product: glucosyltransferase #status predicted <MAN>
 F:1126-1145/Domain: cpl repeat homology <CP1>
 F:1253-1272/Domain: cpl repeat homology <CP2>
 F:1318-1337/Domain: cpl repeat homology <CP3>

Query Match 56.9%; Score 66; DB 2; Length 1375;
 Best Local Similarity 66.7%; Pred. No. 0.1;
 Matches 14; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 1 AINHLSTLEAMSDNDPOYNKD 21
 :|||:|||||:
 DB 507 ANDHLSTLEAMSDNDPPYHLD 527

RESULT 12

AH0630
 4-hydroxyphenylacetate degradation bifunctional isomerase/decarboxylase (EC 5.3.3.-) [im
 C:Species: Salmonella enterica subsp. enterica serovar Typhi
 A>Note: this species has also been called Salmonella typhi
 C:Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov-2002
 C:Accession: AH0630
 R:Parhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher,
 Th. T.; Conerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,
 S.; Moule, S.; O'Gaora, P.
 Nature 413, 848-852, 2001
 A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;
 A>Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serov
 A:Reference number: AB0502; MWID:21534947; PMID:11677608
 A:Accession: AH0630
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-429 <PAR>
 A:Cross-references: GB:AL513382; PIDN:CAD08227.1; PID:g16502275; GSPDB:GN00176
 C:Genetics:
 A:Gene: hpaG
 C:Keywords: intramolecular oxidoreductase; isomerase

Query Match 44.0%; Score 51; DB 2; Length 429;
 Best Local Similarity 47.6%; Pred. No. 5.5;
 Matches 10; Conservative 4; Mismatches 5; Indels 2; Gaps 1;

QY 1 AINHLSTLEAMSD--NDPOYN 19
 :|||:|||||:
 DB 9 AINHRQLDMQAFSGPPYN 29

RESULT 13

T46504
 hypothetical protein DKFZp434M179.1 - human (fragment)
 C:Species: Homo sapiens (man)
 C:Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 04-Feb-2000
 C:Accession: T46504
 R:Ponska, A.; Klein, M.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
 Submitted to the Protein Sequence Database, January 2000
 A:Reference number: Z23029
 A:Accession: T46504
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-632 <AAA>
 A:Cross-references: EMBL:AL137338
 A:Experimental source: adult testis; clone DKFZp434M179
 C:Genetics:
 A>Note: DKFZp434M179.1

Query Match 42.2%; Score 49; DB 2; Length 632;
 Best Local Similarity 52.9%; Pred. No. 17;
 Matches 9; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 5 LSTLEAMSDNDPOYNKD 21
 :|||:|||||:
 DB 119 IMVLGASEPDPYHLD 135

RESULT 14

AZ2493
 hypothetical protein alr7125 [imported] - Nostoc sp. (strain PCC 7120) plasmid pCC7120a1
 C:Species: Nostoc sp. PCC 7120
 A>Note: Nostoc sp. strain PCC 7120 is a synonym of Arabaena sp. strain PCC 7120
 C:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Dec-2002
 C:Accession: AZ2493
 R:Kaneko, T.; Nakamura, Y.; Molk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriiguchi,
 Nakazaki, N.; Shimo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S.
 DNA Res. 8, 205-213, 2001
 A>Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Ana
 A:Reference number: AB1807; MWID:21595285; PMID:11759840
 A:Accession: AZ2493
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-369 <KUR>
 A:Cross-references: GB:BA000020; PIDN:BAF78209.1; PID:g17135663; GSPDB:GN00180
 A:Experimental source: strain PCC 7120
 C:Genetics:
 A:Gene: alr7125
 A:Genome: plasmid

Query Match 40.9%; Score 47.5; DB 2; Length 369;
 Best Local Similarity 40.0%; Pred. No. 16;
 Matches 10; Conservative 3; Mismatches 5; Indels 7; Gaps 1;

QY 4 HSLT-----EAMSDNDPOYNKD 21
 :|||:|||||:
 DB 6 HSLTQGVVWNAWRENNDINPD 30

RESULT 15

C86205
 hypothetical protein [imported] - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
 C:Accession: C86205
 R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
 Chin, C.W.; Chung, M.K.; Com, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
 Anen, N.F.; Hughes, B.; Huizar, L.
 Nature 408, 816-820, 2000
 A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
 C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lures, J.S.; Maiti, R.; Marziani,
 Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
 A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Talion,
 Ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
 A>Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
 A:Reference number: AB6141; MWID:21016719; PMID:11130712
 A:Accession: C86205
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-175 <STO>
 A:Cross-references: GB:AF005172; NID:g8954041; PIDN:AAF82215.1; GSPDB:GN00141
 C:Genetics:
 A:Map position: 1

Query Match 40.5%; Score 47; DB 2; Length 175;
 Best Local Similarity 46.2%; Pred. No. 8.2;
 Matches 6; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

QY 2 INHSLTLEAMSDN 14
 :|||:|||||:
 DB 108 INHQEVYDAMSDH 120

Search completed: November 13, 2003, 09:50:25
 Job time: 14.0379 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 13, 2003, 09:31:40 ; Search time 7.56398 Seconds

(without alignments)
130.561 Million cell updates/sec

Title: US-09-290-049a-11

Perfect score: 116
Sequence: 1 AINHLSTLEAMSNDPOYNKD 21

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 127863 segs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_41:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	116	100.0	1462	1	GTFD_STRMU P49331 streptococc
2	81	69.8	1365	1	GTFG_STRDO P29336 streptococc
3	75	64.7	1476	1	GTFB_STRMU P08987 streptococc
4	74	63.8	1592	1	GTF2_STRDO P27470 streptococc
5	74	63.8	1597	1	GTF1_STRDO P11001 streptococc
6	66	56.9	1455	1	GTRC_STRMU P13470 streptococc
7	51	44.0	429	1	HPAG_SALDU Q9npus s 4-hydroxy
8	49	42.2	759	1	SC63_HUMAN Q9ugp8 homo sapien
9	49	42.2	759	1	SC63_MOUSE Q9vne0 mus musculu
10	47	40.5	418	1	VS12_REOVL P1314 revolutus (t
11	46	39.7	418	1	VS12_REOVL Q8wxk8 homo sapien
12	46	39.7	3060	1	BPEB_HUMAN Q9m6k1 arabidopsis
13	45.5	39.2	560	1	PMG2_ARATH Q97ws6 clostridium
14	45	38.8	644	1	SVN_CICAB Q01377 coynebacte
15	45	38.8	657	1	CSP1_CORGL Q9y719 scytosacch
16	45	38.8	2258	1	WOPD_SCHPO P35157 haemophilus
17	44.5	38.4	241	1	HEB1_HABIN P45991 haemophilus
18	44.5	38.4	241	1	HEB2_HABIN Q971s6 sulfolobus
19	44	37.9	226	1	PYRH_SULTO Q971s6 sulfolobus
20	44	37.9	569	1	SILF_MOUSE Q92805 saccharomy
21	44	37.9	837	1	ROD1_YEAST Q92805 saccharomy
22	44	37.5	825	1	ILAR_HUMAN P24394 homo sapien
23	43	37.1	267	1	COML_NEIGO Q05085 neisseria g
24	43	37.1	267	1	COML_NEIMA Q9jvb7 neisseria m
25	43	37.1	267	1	COML_NEIMA Q9kcb1 neisseria m
26	43	37.1	418	1	VS12_REOVL P03525 revolutus (t
27	43	37.1	443	1	AM2A_ORYSA P27335 oryza sativ
28	43	37.1	443	1	AM2C_ORYSA P27341 oryza sativ
29	43	37.1	480	1	DISA_TRIGA P15503 trimeresura
30	43	37.1	6885	1	SNE2_HUMAN Q8wxh0 homo sapien
31	42.5	36.6	451	1	MYB8_DICDI P34127 dicystoteli
32	42.5	36.6	661	1	PSAB_PROHO P58387 psychotroch
33	42	36.2	405	1	HPCE_EROLI P37352 e homoproteo

34	42	36.2	429	1	HPAG_ECOLI Q46978 e 4-hydroxy
35	42	36.2	447	1	GNT1_RAT Q03325 rattus norv
36	42	36.2	461	1	TR1A_RAT P22934 rattus norv
37	42	36.2	989	1	RPOC_LEUME P44902 leucostoc
38	42	36.2	1097	1	LIFR_HUMAN P43702 homo sapien
39	41.5	35.8	511	1	U171_CAEL Q20063 caenorhabdi
40	41.5	35.8	1616	1	VAED_SCHPO Q09853 schizosacch
41	41	35.3	305	1	SGAU_MYCPN P75294 mycoplasma
42	41	35.3	438	1	GATD_PYRPU Q8u0x0 pyrococcus
43	41	35.3	576	1	CBPS_YEAST P27614 saccharomy
44	41	35.3	591	1	CALX_MOUSE P35564 mus musculu
45	41	35.3	591	1	CALX_RAT P35565 rattus norv

ALIGNMENTS

RESULT 1

ID GTFD_STRMU STANDARD; PRT; 1462 AA.
AC P49331; O69383; O69386; O69389; O69392; O69398;
DT 01-FEB-1996 (Rel. 33, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Glucosyltransferase-S precursor (EC 2.4.1.5) (GTF-S) (Dextranucrase)
GN GTFD OR SMU.910.
OS Streptococcus mutans.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1309;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=GS-5;
RX MEDLINE=91100958; PubMed=2148600;
RA Honda O., Kato C., Kuramitsu H.K.;
RT "Nucleotide sequence of the Streptococcus mutans gtfD gene encoding the glucosyltransferase-S enzyme.";
RL J. Gen. Microbiol. 136:2099-2105(1990).
[2]
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=MT4239 / Serotype C, MT4245 / Serotype B, MT4251 / Serotype F, MT4467 / Serotype E, and MT8148 / Serotype C;
RX MEDLINE=96231643; PubMed=9570124;
RA Fujiwara T., Terao Y., Hoshino T., Kawabata S., Ooshima T., Sobue S., Kimura S., Hamada S.;
RT "Molecular analyses of glucosyltransferase genes among strains of Streptococcus mutans.";
RL FEMS Microbiol. Lett. 161:331-336(1998).
[3]
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=UA159 / Serotype C; ATCC 700610 / Serotype C;
RX MEDLINE=22295063; PubMed=12397186;
RA Ajdic D., McShan W.M., McLaughlin R.E., Savic G., Chang J., Carson M.B., Primeaux C., Tian R., Kenton S., Jia H., Lin S., Qian Y., Li S., Zhu H., Najjar F., Lai H., White J., Roe B.A., Ferretti J.J.;
RT "Genome sequence of Streptococcus mutans UA159, a cariogenic dental pathogen.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:14434-14439(2002).
CC -!- FUNCTION: PRODUCTION OF EXTRACELLULAR GLUCANS, THAT ARE THOUGHT TO PLAY A KEY ROLE IN THE DEVELOPMENT OF THE DENTAL PLAQUE BECAUSE OF THEIR ABILITY TO ADHERE TO SMOOTH SURFACES AND MEDATE THE AGGREGATION OF BACTERIAL CELLS AND FOOD DEBRIS.
CC -!- CATALYTIC ACTIVITY: Sucrose + ((1,6)-alpha-D-glucosyl) (N) = D-fructose + ((1,6)-alpha-D-glucosyl) (N+1).
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- DISEASE: DENTAL CARIES.
CC -!- MISCELLANEOUS: GTF-I SYNTHESIZES WATER-INSOLUBLE GLUCANS (ALPHA 1,3-LINKED GLUCOSE AND SOME 1,6 LINKAGES), GTF-S SYNTHESIZES WATER-SOLUBLE GLUCANS (ALPHA 1,6-GLUCOSE), GTF-SI SYNTHESIZES BOTH FORMS OF GLUCANS.
CC -!- SIMILARITY: TO OTHER GLUCOSYLTRANSFERASES AND SOME TO A GLUCAN-BINDING PROTEIN FROM S. MUTANS.

CC -1- SIMILARITY: Contains 6 cell wall binding repeats.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; M29296; AAA26895.1; -
 CC EMBL; D88653; BAA26103.1; -
 CC EMBL; D88656; BAA26107.1; -
 CC EMBL; D88659; BAA26111.1; -
 CC EMBL; D88662; BAA26115.1; -
 CC EMBL; D89379; BAA26121.1; -
 CC EMBL; AE014932; AAN58619.1; -
 CC InterPro; IPR002479; CW_binding.
 CC InterPro; IPR003318; Glyco_hydro_70.
 CC Pfam; PF02324; Glyco_hydro_70; 1.
 CC Transferase; Glycosyltransferase; Signal; Repeat; Dental caries;
 CC Complete proteome.
 CC KW
 CC Complete proteome.
 CC FT SIGNAL 1 ?
 CC FT CHAIN 1 1462
 CC FT DOMAIN 1232 1423
 CC FT REPEAT 1232 1295
 CC FT REPEAT 1296 1359
 CC FT REPEAT 1360 1423
 CC FT REPEAT 10 10
 CC FT VARIANT 19 19
 CC FT VARIANT 58 58
 CC FT VARIANT 68 68
 CC FT VARIANT 81 81
 CC FT VARIANT 113 113
 CC FT VARIANT 122 122
 CC FT VARIANT 132 132
 CC FT VARIANT 135 135
 CC FT VARIANT 137 137
 CC FT VARIANT 202 202
 CC FT VARIANT 255 255
 CC FT VARIANT 275 275
 CC FT VARIANT 288 288
 CC FT VARIANT 301 301
 CC FT VARIANT 313 313
 CC FT VARIANT 317 317
 CC FT VARIANT 328 328
 CC FT VARIANT 350 350
 CC FT VARIANT 628 628
 CC FT VARIANT 688 688
 CC FT VARIANT 726 732
 CC FT VARIANT 726 730
 CC FT VARIANT 762 762
 CC FT VARIANT 964 964
 CC FT VARIANT 1019 1019
 CC FT VARIANT 1060 1060
 CC FT VARIANT 1060 1060
 CC FT VARIANT 1080 1080
 CC FT VARIANT 1142 1142
 CC FT VARIANT 1142 1142
 CC FT VARIANT 1198 1198
 CC FT VARIANT 1220 1220
 CC FT VARIANT 1280 1280
 CC FT VARIANT 1282 1282
 CC FT VARIANT 1290 1290
 CC FT VARIANT 1311 1311
 CC FT VARIANT 1311 1311

FT VARIANT 1403 1403 G -> D (IN STRAINS GS-5 AND MT4467).
 FT VARIANT 1425 1425 G -> R (IN STRAIN GS-5).
 FT VARIANT 1449 1449 R -> K (IN STRAIN MT4467).
 FT CONFLICT 1428 1462 RYDKNSGNMVMYVTLANGRRIGIDRWGIARY -> VY
 FT R (IN REF. 1).
 SQ SEQUENCE 1462 AA; 163387 MW; CE4A279C4D708645 CRC64;
 Query Match 100.0%; Score 116; DB 1; Length 1462;
 Best Local Similarity 100.0%; Pred. No. 9e-10;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 AINHLSTLEAMSDNDPOYNKD 21
 Db 495 AINHLSTLEAMSDNDPOYNKD 515
 RESULT 2
 GTF5_STRDO STANDARD; PRT; 1365 AA.
 ID GTF5_STRDO
 AC P29336;
 DT 01-DEC-1992 (Rel. 24, Created)
 DT 01-DEC-1992 (Rel. 24, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Glucosyltransferase-5 precursor (EC 2.4.1.5) (GTF-S) (Dextranucrase)
 DE (Sucrose 6-glucosyltransferase).
 GN GTF5.
 OS Streptococcus downei (Streptococcus sobrinus).
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 OK NCBI_TaxId=1317;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MFE28;
 RX MEDLINE=90316665; PubMed=2142479;
 RA Gilmore K.S., Russell R.R., Ferretti J.J.;
 RT Analysis of the Streptococcus downei gtf5 gene, which specifies a
 RT glucosyltransferase that synthesizes soluble glucans.";
 RL Infect. Immun. 58:2452-2458(1990).
 CC -1- FUNCTION: PRODUCTION OF EXTRACELLULAR GLUCANS, THAT ARE THOUGHT
 CC TO PLAY A KEY ROLE IN THE DEVELOPMENT OF THE DENTAL PLAQUE BECAUSE
 CC OF THEIR ABILITY TO ADHERE TO SMOOTH SURFACES AND MEDATE THE
 CC AGGREGATION OF BACTERIAL CELLS AND FOOD DEBRIS.
 CC -1- CATALYTIC ACTIVITY: Sucrose + [(1,6)-alpha-D-glucosyl] (N+1) = D-
 CC fructose + [(1,6)-alpha-D-glucosyl] (N+1).
 CC -1- ENZYME REGULATION: GLUCAN SYNTHESIS BY GTF-S IS INDEPENDENT OF
 CC PRIMER GLUCAN UNLIKE GTF-1.
 CC -1- DISEASE: DENTAL CARIES.
 CC -1- MISCELLANEOUS: GTF-S SYNTHESIZES WATER-SOLUBLE GLUCANS (ALPHA
 CC 1,6-GLUCOSE).
 CC -1- SIMILARITY: TO OTHER GLUCOSYLTRANSFERASES AND SOME TO A GLUCAN-
 CC BINDING PROTEIN FROM S.MUTANS.
 CC -1- SIMILARITY: Contains 10 cell wall binding repeats.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; M30943; AAA26898.1; -
 CC InterPro; IPR002479; CW_binding.
 CC InterPro; IPR003318; Glyco_hydro_70.
 CC Pfam; PF01473; CW_binding_1; 8.
 CC Pfam; PF02324; Glyco_hydro_70; 1.
 CC Transferase; Glycosyltransferase; Signal; Repeat; Dental caries.
 CC KW
 CC SIGNAL 1 36
 CC CHAIN 37 1365
 CC REPEAT 157 177
 CC REPEAT 178 197
 CC DOMAIN 198 1061
 CC REPEAT 1062 1082
 CC OR 37 (POTENTIAL).
 CC GLUCOSYLTRANSFERASE-S.
 CC CELL WALL BINDING 1.
 CC CATALYTIC (APPROXIMATE).
 CC CELL WALL BINDING 3.

	Query Match	Best Local Similarity	Score 81;	DB 1;	Length 1365;
	Matches 16;	Conservative 1;	Pred. No. 0.00021;	Mismatches 4;	Indels 0; Gaps 0
0y	1				
Db	467	ALDH3L5LEAMSNDNDYXKD 467			
RESULT 3					
ID	GF6F_STRMU	STANDARD;	PRT: 1476 AA.		
AC	089987; 069381; 069384; 069387; 069390; 069396;				
DT	28-FEB-2003 (Rel. 41, last sequence update)				
DT	15-SEP-2003 (Rel. 42, last annotation update)				
DE	Glucosyltransferase-I precursor (EC 2.4.1.5) (GTF-II) (Dextranucrase)				
DE	(Sucrose 6-glucosyltransferase).				
GN	GF6F OR SMU.1004.				
OS	Streptococcus mutans.				
OC	Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;				
OC	Streptococcus.				
OX	NCBI_TaxID=1309;				
OX	[1]				
RN	SEQUENCE FROM N.A.				
RP	STRAIN=GS-5;				
RC	MEDLINE=87308013; PubMed=3040685;				
RA	Shiroza T., Ueda S., Kuramitsu H.K.;				
RT	"Sequence analysis of the gtfB gene from Streptococcus mutans.";				
RL	J. Bacteriol. 169:4263-4270(1987).				
RN	[2]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=MT4239 / Serotype C, MT4245 / Serotype E, MT4251 / Serotype F,				
RC	MT4467 / Serotype E, and MT8148 / Serotype C;				
RC	MEDLINE=98231643; PubMed=9570124;				
RA	Fujiwara T., Terao Y., Hoshino T., Kawabata S., Ooshima T., Sobue S.,				
RA	Kimura S., Hamada S.;				
RT	"Molecular analyses of glucosyltransferase genes among strains of				
RT	Streptococcus mutans.";				
RL	FEMS Microbiol. Lett. 161:331-336(1998).				
RN	[3]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=UAI59 / ATCC 700610 / Serotype C;				
RC	MEDLINE=22295063; PubMed=12397186;				
RA	Ajdic D., McGhan W.M., McLaughlin R.E., Savic G., Chang J.;				
RA	Carson M.B., Primeaux C., Tan R., Kenton S., Jia H., Lin S., Qian Y.,				
RA	Li S., Zhu H., Najjar F., Lai H., White J., Roe B.A., Ferretti J.D.;				
RT	"Genome sequence of Streptococcus mutans UAI59, a cariogenic dental				
RT	pathogen.";				
RL	Proc. Natl. Acad. Sci. U.S.A. 99:14434-14439(2002).				
CC	-1- FUNCTION: PRODUCTION OF EXTRACELLULAR GLUCANS, THAT ARE THOUGHT				
CC	TO PLAY A KEY ROLE IN THE DEVELOPMENT OF THE DENTAL PLAQUE BECAUSE				
CC	OF THEIR ABILITY TO ADHERE TO SMOOTH SURFACES AND MEDIANE THE				
CC	AGGREGATION OF BACTERIAL CELLS AND FOOD DEBRIS.				
CC	-1- CATALYTIC ACTIVITY: Sucrose + ((1,6)-alpha-D-glucosyl) (N) = D-				
CC	fructose + ((1,6)-alpha-D-glucosyl) (N+1).				
CC	-1- SUBCELLULAR LOCATION: Secreted.				
CC	-1- DISEASE: DENTAL CARIES.				
CC	-1- MISCELLANEOUS: GTF-I SYNTHESIZES WATER-INSOLUBLE GLUCANS (ALPHA				
CC	1,3-LINKED GLUCOSE AND SOME 1,6 LINKAGES), GTF-S SYNTHESIZES				
CC	WATER-SOLUBLE GLUCANS (ALPHA 1,6-GLUCOSE). GTF-SI SYNTHESIZES BOTH				
CC	FORMS OF GLUCANS				
CC	-1- SIMILARITY: TO OTHER GLUCOSYLTRANSFERASES AND SOME TO A GLUCAN-				
CC	BINDING PROTEIN FROM S. MUTANS				

CC	-1- SIMILARITY: Contains 10 cell wall binding repeats.
CC	
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration
CC	between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC	the European Bioinformatics Institute. There are no restrictions on its
CC	use by non-profit institutions as long as its content is in no way
CC	modified and this statement is not removed. Usage by and for commercial
CC	entities requires a license agreement (See http://www.isb.ch/announce/
CC	or send an email to license@isb-sib.ch).
CC	
DR	EMBL; M17361; AAA88588.1; -
DR	EMBL; D88651; BAA26101.1; -
DR	EMBL; D88654; BAA26105.1; -
DR	EMBL; D88657; BAA26109.1; -
DR	EMBL; D88660; BAA26113.1; -
DR	EMBL; D89377; BAA26119.1; -
DR	EMBL; AE014940; AAN58705.1; -
DR	InterPro; IPR002479; CW binding.
DR	InterPro; IPR003318; Glyco_hydro_70.
DR	Pfam; PF01473; CW_binding_1; 1.
KW	Pfam; PF02324; Glyco_hydro_70; 1.
KW	Transferase; Glycosyltransferase; Signal; Repeat; Dental caries;
KW	Complete proteome.
FT	SIGNAL 1 34
FT	CHAIN 35 1476
FT	DOMAIN 35 1051
FT	DOMAIN 1097 1476
FT	REPEAT 1097 1130
FT	DOMAIN 1161 1470
FT	REPEAT 1161 1210
FT	REPEAT 1225 1275
FT	REPEAT 1290 1340
FT	REPEAT 1355 1405
FT	REPEAT 1420 1470
FT	VARIANT 62 62
FT	VARIANT 65 65
FT	VARIANT 68 68
FT	VARIANT 78 78
FT	VARIANT 86 86
FT	VARIANT 89 89
FT	VARIANT 168 168
FT	VARIANT 276 276
FT	VARIANT 399 399
FT	VARIANT 474 474
FT	VARIANT 512 512
FT	VARIANT 519 519
FT	VARIANT 701 701
FT	VARIANT 708 708
FT	VARIANT 938 938
FT	VARIANT 952 957
FT	VARIANT 963 964
FT	VARIANT 968 970
FT	VARIANT 1086 1086
FT	VARIANT 1158 1158
FT	VARIANT 1163 1163
FT	VARIANT 1168 1168
FT	VARIANT 1182 1182
FT	VARIANT 1234 1234
FT	VARIANT 1263 1263
FT	VARIANT 1264 1264
FT	VARIANT 1272 1272
FT	VARIANT 1329 1329
FT	VARIANT 1394 1394
FT	
CC	
CC	POTENTIAL.
CC	GLUCOSYLTRANSFERASE-I.
CC	CATALYTIC (APPROXIMATE).
CC	GLUCAN-BINDING (APPROXIMATE).
CC	A REPEAT.
CC	5 X TANDEM REPEATS.
CC	1.
CC	2.
CC	3.
CC	4.
CC	5.
CC	S -> T (IN STRAIN MT4239).
CC	T -> I (IN STRAIN GS-5).
CC	V -> A (IN STRAINS GS-5, MT4245, MT4251,
CC	MT4467 AND MT8148).
CC	Q -> P (IN STRAIN MT4251).
CC	I -> S (IN STRAINS GS-5, MT4245, MT4251,
CC	MT4467 AND MT8148).
CC	S -> F (IN STRAIN MT4251).
CC	K -> N (IN STRAIN MT4251).
CC	S -> D (IN STRAINS GS-5, MT4467 AND
CC	MT8148).
CC	N -> R (IN STRAIN MT4239).
CC	I -> T (IN STRAIN MT4239).
CC	K -> R (IN STRAIN MT8148).
CC	F -> I (IN STRAIN MT8148).
CC	T -> I (IN STRAIN MT8148).
CC	A -> V (IN STRAIN MT8148).
CC	F -> L (IN STRAIN MT8148).
CC	FGKPEV -> YGTPVA (IN STRAINS GS-5, MT4239
CC	AND MT4467).
CC	SV -> NT (IN STRAINS GS-5, MT4239 AND
CC	MT4467).
CC	ADG -> VDQ (IN STRAINS GS-5, MT4239 AND
CC	MT4467).
CC	A -> T (IN STRAIN MT4239).
CC	S -> N (IN STRAIN MT4239).
CC	H -> Y (IN STRAIN MT4251).
CC	E -> K (IN STRAIN MT8148).
CC	Y -> C (IN STRAIN MT8148).
CC	A -> P (IN STRAIN MT4239).
CC	R -> P (IN STRAIN MT8148).
CC	Y -> H (IN STRAINS GS-5 AND MT4467).
CC	R -> H (IN STRAINS GS-5, MT4239, MT4467
CC	AND MT8148).
CC	S -> G (IN STRAINS GS-5, MT4239, MT4467
CC	AND MT8148).
CC	H -> Y (IN STRAINS GS-5 AND MT4467).
CC	Y -> H (IN STRAINS GS-5, MT4239, MT4467
CC	AND MT8148).

```

FT  VARIANT      1402      1402      S -> G (IN STRAINS GS-5, MT4239, MT4467
FT  VARIANT      1459      1459      AND MT8148).
FT  VARIANT      570      570      Y -> H (IN STRAIN MT4467).
FT  CONFLICT     800      817      R -> A (IN REF. 1).
FT  CONFLICT     800      817      ADQDVRAASTAPSTDGK -> LIMEFLRLARPHQMA
FT  CONFLICT     1310     1310      (IN REF. 1).
FT  CONFLICT     1476     165846      H -> L (IN REF. 1).
SQ  SEQUENCE     1476 AA; 165846 MW; 9C6E09F731B4C6CF CRC64;

Query Match      64.7%; Score 75; DB 1; Length 1476;
Best Local Similarity 71.4%; Pred. No. 0.002;
Matches 15; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

Qy  1 ANHSTIEAWSNDPOYNKD 21
Db  481 ANHSTIEAWSNDPTIYHD 501

RESULT 4
GTF2_STRDO
ID  GTF2_STRDO      STANDARD;      PRT; 1592 AA.
AC  P27470;
DT  01-AUG-1992 (Rel. 23, Created)
DT  01-AUG-1992 (Rel. 23, Last sequence update)
DT  15-SEP-2003 (Rel. 42, Last annotation update)
DE  Glucosyltransferase-1 precursor (EC 2.4.1.5) (GTF-1) (Dextranucrase)
DE  (Sucrose 6-glucosyltransferase).
OS  Streptococcus downei (Streptococcus sobrinus).
OC  Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC  Streptococcus.
OX  NCBI_TaxID=1317;
RN  [1]
RP  SEQUENCE FROM N.A.
RX  STRAIN=6715 / Serotype G;
RX  MEDLINE=91123227; PubMed=1704006;
RA  Abo H., Matsumura T., Kodama T., Ohta H., Fukui K., Kato K.,
RA  Kagawa H.;
RT  "Peptide sequences for sucrose splitting and glucan binding within
RT  Streptococcus sobrinus glucosyltransferase (water-insoluble glucan
RT  synthetase).";
RL  J. Bacteriol. 173:989-996(1991).
CC  -1- FUNCTION: PRODUCTION OF EXTRACELLULAR GLUCANS, THAT ARE THOUGHT
CC  TO PLAY A KEY ROLE IN THE DEVELOPMENT OF THE DENTAL PLAQUE BECAUSE
CC  OF THEIR ABILITY TO ADHERE TO SMOOTH SURFACES AND MEDIATE THE
CC  AGGREGATION OF BACTERIAL CELLS AND FOOD DEBRIS.
CC  -1- CATALYTIC ACTIVITY: Sucrose + {(1,6)-alpha-D-glucosyl}(N) = D-
CC  fructose + {(1,6)-alpha-D-glucosyl}(N+1).
CC  -1- SUBCELLULAR LOCATION: Secreted.
CC  -1- DISEASE: DENTAL CARIES.
CC  -1- MISCELLANEOUS: GTF-1 SYNTHESIZES WATER-INSOLUBLE GLUCANS (ALPHA
CC  1,3-LINKED GLUCOSE AND SOME 1,6 LINKAGES), GTF-S SYNTHESIZES
CC  WATER-SOLUBLE GLUCANS (ALPHA 1,6-GLUCOSE). GTF-SI SYNTHESIZES BOTH
CC  FORMS OF GLUCANS.
CC  -1- SIMILARITY: TO OTHER GLUCOSYLTRANSFERASES AND SOME TO A GLUCAN-
CC  BINDING PROTEIN FROM S.MUTANS.
CC  -1- SIMILARITY: Contains 16 cell wall binding repeats.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
CC  EMBL; D90213; BAAL4241.1;
CC  InterPro; IPR002479; CM binding.
CC  InterPro; IPR003318; Glyco_hydro_70.
CC  Pfam; PF01473; CM binding_1; 13.
CC  Pfam; PF02324; Glyco_hydro_70; 1.
CC  Transferrase; Glycosyltransferase; Signal; Repeat; Dental caries.
CC  SIGNAL 1
CC  CHAIN 39 1592 GLUCOSYLTRANSFERASE-I.

```

```

FT  DOMAIN      39      1044      CATALYTIC (APPROXIMATE).
FT  DOMAIN      1093     1592      GLUCAN-BINDING (APPROXIMATE).
FT  DOMAIN      1093     1592      6.5 X TANDEM REPEATS.
FT  REPEAT      1093     1142      1.
FT  REPEAT      1158     1207      2.
FT  REPEAT      1222     1272      3.
FT  REPEAT      1287     1337      4.
FT  REPEAT      1402     1451      5.
FT  REPEAT      1514     1563      6.
FT  REPEAT      1577     1592      7 (INCOMPLETE).
SQ  SEQUENCE     1592 AA; 176167 MW; BCOA66D079351ECF CRC64;

Query Match      63.8%; Score 74; DB 1; Length 1592;
Best Local Similarity 66.7%; Pred. No. 0.0031;
Matches 14; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Qy  1 ANHSTIEAWSNDPOYNKD 21
Db  477 ANHSTIEAWSNDPTIYHD 497

RESULT 5
GTF1_STRDO
ID  GTF1_STRDO      STANDARD;      PRT; 1597 AA.
AC  P11001;
DT  01-JUL-1989 (Rel. 11, Created)
DT  01-JUL-1989 (Rel. 11, Last sequence update)
DT  15-SEP-2003 (Rel. 42, Last annotation update)
DE  Glucosyltransferase-1 precursor (EC 2.4.1.5) (GTF-1) (Dextranucrase)
DE  (Sucrose 6-glucosyltransferase).
OS  Streptococcus downei (Streptococcus sobrinus).
OC  Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC  Streptococcus.
OX  NCBI_TaxID=1317;
RN  [1]
RP  SEQUENCE FROM N.A.
RX  STRAIN=MFE28;
RX  MEDLINE=87308014; PubMed=3040686;
RA  Ferretti J.J., Gilpin M.L., Russell R.R.B.;
RA  "Nucleotide sequence of a glucosyltransferase gene from Streptococcus
RA  sobrinus MFE28.";
RL  J. Bacteriol. 165:4271-4278(1987).
CC  -1- FUNCTION: PRODUCTION OF EXTRACELLULAR GLUCANS, THAT ARE THOUGHT
CC  TO PLAY A KEY ROLE IN THE DEVELOPMENT OF THE DENTAL PLAQUE BECAUSE
CC  OF THEIR ABILITY TO ADHERE TO SMOOTH SURFACES AND MEDIATE THE
CC  AGGREGATION OF BACTERIAL CELLS AND FOOD DEBRIS.
CC  -1- CATALYTIC ACTIVITY: Sucrose + {(1,6)-alpha-D-glucosyl}(N) = D-
CC  fructose + {(1,6)-alpha-D-glucosyl}(N+1).
CC  -1- SUBCELLULAR LOCATION: Secreted.
CC  -1- DISEASE: DENTAL CARIES.
CC  -1- MISCELLANEOUS: GTF-1 SYNTHESIZES WATER-INSOLUBLE GLUCANS (ALPHA
CC  1,3-LINKED GLUCOSE AND SOME 1,6 LINKAGES), GTF-S SYNTHESIZES
CC  WATER-SOLUBLE GLUCANS (ALPHA 1,6-GLUCOSE). GTF-SI SYNTHESIZES BOTH
CC  FORMS OF GLUCANS.
CC  -1- SIMILARITY: TO OTHER GLUCOSYLTRANSFERASES AND SOME TO A GLUCAN-
CC  BINDING PROTEIN FROM S.MUTANS.
CC  -1- SIMILARITY: Contains 19 cell wall binding repeats.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
CC  EMBL; M17391; AAC63063.1;
CC  InterPro; IPR002479; CM binding.
CC  InterPro; IPR003318; Glyco_hydro_70.
CC  Pfam; PF01473; CM binding_1; 16.
CC  Pfam; PF02324; Glyco_hydro_70; 1.
CC  Transferase; Glycosyltransferase; Signal; Repeat; Dental caries.
CC  SIGNAL 1
CC  CHAIN 39 1592 GLUCOSYLTRANSFERASE-I.

```


Query Match
Best Local Similarity 56.9%; Score 66; DB 1; Length 1455;
Pred. No. 0.047;
Matches 14; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

1 AINHLSTLEAWSDNDPOYNKD 21
507 ANDHLSTLEAWSYNDPTIYLD 527

RESULT 7
HPAG_SALDU STANDARD; PRT; 429 AA.
ID HPAG_SALDU
AC Q9RP05; 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE 4-hydroxyphenylacetate degradation bifunctional
isomerase/decarboxylase [includes: 2-hydroxyhepta-2,4-diene-1,7-dioate
isomerase (EC 5.3.3.-) (HHD isomerase); 5-carboxymethyl-2-oxo-hex-3-
ene-1,7-dioate decarboxylase (EC 4.1.1.-) (OPE decarboxylase)].
GN HPAG.
OS Salmonella dublin.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Salmonella.
OX NCBI_TaxID=98360;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=2229;
RA Galinov E.E., Wood M.W., Hedges S.;
RT "Characterization of the hpa genetic locus from Salmonella dublin.";
RT Submitted (Apr-1999) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: DECARBOXYLATES OPEPT (5-OXO-PENT-3-ENE-1,7,5-
TRICARBOXYLIC ACID) INTO HHD (2-HYDROXY-HEPT-2,4-DIENE-1,7-
DIOATE) AND ISOMERIZES IT TO OHD (2-OXO-HEPT-3-ENE-1,7-DIOATE).
CC -1- PATHWAY: 4-hydroxyphenylacetate catabolism.
CC -1- SIMILARITY: Belongs to the FAH family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL: AF144422; AAD53501.1; -.
DR InterPro: IPR002529; FAA hydrolase.
DR Pfam: PF01557; FAA hydrolase; 2.
KM Aromatic hydrocarbons catabolism; Lyase; Isomerase; Repeat;
KW Multifunctional enzyme.
FT REPEAT 1 215 APPROXIMATE.
FT REPEAT 216 429 APPROXIMATE.
SQ SEQUENCE 429 AA; 47185 MW; D32723999AC4084E CRC64;

Query Match
Best Local Similarity 44.0%; Score 51; DB 1; Length 429;
Pred. No. 2.3;
Matches 10; Conservative 4; Mismatches 5; Indels 2; Gaps 1;

1 AINHLSTLEAWSD--NDPOYN 19
9 ALNHRSQLDAMQEAFAFSQPPYN 29

RESULT 8
SC63 HUMAN STANDARD; PRT; 759 AA.
ID SC63 HUMAN
AC Q9UGP8; Q95380; Q9NTE0;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Translocation protein SC63 homolog.
GN SEC63 OR SEC63L.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20008798; PubMed=10543453;
RA Skowronek M.H., Rotter M., Haas I.G.;
RT "Molecular characterization of a novel mammalian DnaJ-like Sec63p
homolog.";
RT Biol. Chem. 380:1133-1138(1999).
RN [2]
RP SEQUENCE FROM N.A.
RA Hartmann E., Prehn S.;
RT Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE OF 128-759 FROM N.A.
RC TISSUE=Testis;
RA Poulik A., Klein M., Nemes H.-W., Gassenhuber J., Wiemann S.;
RT Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: Required for preprotein translocation.
CC -1- SUBUNIT: COMPLEX THAT CONTAINS SEC61, SEC62 AND SEC63.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Endoplasmic
CC reticulum.
CC -1- SIMILARITY: Contains 1 J domain.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL: AF011779; CAB46275.1; -.
DR EMBL: AF100144; AAC83375.1; -.
DR EMBL: AL137338; CAB70701.1; -.
DR PIR: T46504; T46504.
DR HSSP: P25685; 1HDJ.
DR Genew: HGNC:21082; SEC63.
DR InterPro: IPR001623; DnaJ N.
DR InterPro: IPR003095; Hsp_DnaJ.
DR InterPro: IPR004179; Sec63.
DR Pfam: PF00226; DnaJ; 1.
DR Pfam: PF02889; Sec63; 1.
DR PRINTS: PR00625; DNAJPROTEIN.
DR SMART: SM00271; DnaJ; 1.
DR PROSITE: PS00636; DnaJ_1; FALSE_NEG.
DR PROSITE: PS50076; DnaJ_2; 1.
KM Chapterone; Transmembrane; Protein transport; Endoplasmic reticulum;
KW Microsome.
FT INIT MET
FT DOMAIN 1 13 BY SIMILARITY.
FT TRANSMEM 14 13 LUMENAL (POTENTIAL).
FT DOMAIN 35 68 POTENTIAL.
FT DOMAIN 35 68 CYTOPLASMIC (POTENTIAL).

```

CC FT TRANSMEM 69 89 POTENTIAL.
CC FT DOMAIN 90 187 LUMENAL (POTENTIAL).
CC FT TRANSMEM 188 208 POTENTIAL.
CC FT DOMAIN 209 759 CYTOPLASMIC (POTENTIAL).
CC FT DOMAIN 103 164 J-DOMAIN.
CC FT DOMAIN 733 759 ASP/GLU-RICH (HIGHLY ACIDIC).
CC SQ SEQUENCE 759 AA; 87865 MW; BA13AB5B2C53F08 CRC64;

Query Match 42.2%; Score 49; DB 1; Length 759;
Best Local Similarity 52.9%; Pred. No. 9.3;
Matches 9; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

Oy 5 LSI:EWASNDPQYKND 21
Db 246 IMVLAGASEPDPQYKND 262

RESULT 9
SC63 MOUSE STANDARD; PRT: 759 AA.
AC Q8VHE0; Q8VHE3; Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Translocation protein SEC63 homolog.
CN SEC63 OR SEC63L.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_Taxid=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J;
RA Mack M., Noben-Trauth K.;
RT "Phenotypic characterization and positional mapping of the mouse
RT deafness mutation jackson circler (jc).";
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Scheffer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares W.B., Bonaldo M.P., Casavant T.L., Scheetz T.R.,
RA Brownstein M.J., Ustin T.B., Toshynski S., Carrinci P., Prange C.,
RA Raha S.S., Loguercio N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Buterfield V.S.N., Krzywicki M.J., Skalska U., Smalins D.E.,
RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences.";
CC Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -1- FUNCTION: Required for preproteins translocation (By similarity).
CC -1- SUBUNIT: COMPLEX THAT CONTAINS SEC61, SEC62 AND SEC63 (By
CC similarity).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Endoplasmic
CC reticulum.
CC -1- SIMILARITY: Contains 1 J domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/

```

```

CC or send an email to license@isb-sib.ch).
CC EMBL: AY024346; AAK00580.1; -
CC EMBL: BC019366; AAH19366.1; -
CC EMBL: BC031846; AAH31846.1; -
CC WGI: 2155302; Sec63.
CC InterPro: IPR001623; DnaJ_N.
CC InterPro: IPR003095; Hsp DnaJ.
CC InterPro: IPR004179; Sec63.
CC Pfam: PF00226; DnaJ; 1.
CC Pfam: PF02889; Sec63; 1.
CC PRINTS: PR00625; DNAJPROTEIN.
CC SMART: SM00271; DnaJ; 1.
CC SMART: SM00611; SEC63; 1.
CC PROSITE: PS00636; DNAJ_1; FALSE_NEG.
CC PROSITE: PS50076; DNAJ_2; 1.
CC Chapterone; Transmembrane; Protein transport; Endoplasmic reticulum;
CC KW Microsome.
CC FT INIT MET 0 0 BY SIMILARITY.
CC FT DOMAIN 1 13 LUMENAL (POTENTIAL).
CC FT TRANSMEM 14 34 POTENTIAL.
CC FT DOMAIN 35 68 CYTOPLASMIC (POTENTIAL).
CC FT TRANSMEM 69 89 POTENTIAL.
CC FT DOMAIN 90 187 LUMENAL (POTENTIAL).
CC FT TRANSMEM 188 208 POTENTIAL.
CC FT TRANSMEM 209 759 CYTOPLASMIC (POTENTIAL).
CC FT DOMAIN 103 164 J-DOMAIN.
CC FT DOMAIN 733 759 ASP/GLU-RICH (HIGHLY ACIDIC).
CC FT CONFLICT 47 47 N -> S (IN REF. 1).
CC FT CONFLICT 266 266 K -> R (IN REF. 1).
CC SQ SEQUENCE 759 AA; 87710 MW; D1406C01F07232DA CRC64;

Query Match 42.2%; Score 49; DB 1; Length 759;
Best Local Similarity 52.9%; Pred. No. 9.3;
Matches 9; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

Oy 5 LSI:EWASNDPQYKND 21
Db 246 IMVLAGASEPDPQYKND 262

RESULT 10
VS12_REV01 STANDARD; PRT: 418 AA.
ID VS12_REV01
AC P11314;
DT 01-JUL-1989 (Rel. 11, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 01-OCT-1994 (Rel. 30, Last annotation update)
DE Sigma 2 protein (Core protein).
GN S2.
OS Reovirus (type 1 / strain lang).
OC Viruses; dsRNA viruses; Reoviridae; Orthoreovirus.
OX NCBI_Taxid=10884;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92015462; PubMed=1920614;
RA Dermody T.S., Schiff L.A., Nibert M.L., Coombs K.M., Fields B.N.;
RT "The S2 gene nucleotide sequences of prototype strains of the three
RT reovirus serotypes: characterization of reovirus core protein sigma
RT 2.";
CC J. Virol. 65:5721-5731(1991).
CC [2]
CC SEQUENCE FROM N.A.
RX MEDLINE=88024195; PubMed=3663211;
RA George C.X., Crowe A., Munemitsu S.M., Atwater J.A., Samuel C.E.;
RT "Biosynthesis of reovirus-specified polypeptides. Molecular cDNA
RT cloning and nucleotide sequence of the reovirus serotype 1 lang
RT strain s2 mRNA which encodes the virion core polypeptide sigma 2.";
CC Biochem. Biophys. Res. Commun. 147:1153-1161(1987).
CC -1- CAUTION: REF. 2 SEQUENCE DIFFERS FROM THAT SHOWN FROM POSITION 323
CC ONWARD AND IS SHORTER (135 AA) DUE TO A FRAMESHIFT.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

```

```
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; L19774; AAA47239.1; -.
DR EMBL; M17598; AAA47278.1; ALT_FRAME.
DR PIR; A41306; FOXRL2.
DR InterPro; IPR004317; Sigma_1_2.
DR Pfam; PF03084; Sigma_1_2; 1.
DR ProDom; PD004419; Sigma_1_2; 1.
KM Core protein.
FT CONFLICT 89 89 I -> V (IN REF. 2).
FT CONFLICT 205 205 H -> Y (IN REF. 2).
SQ SEQUENCE 418 AA; 47110 MW; 4B30CF9217EA1DDB CRC64;

Query Match 40.5%; Score 47; DB 1; Length 418;
Best Local Similarity 61.5%; Pred. No. 9.4;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 8 LEAWSNDPOYNK 20
|||::|||::
Db 372 LEAWARDDOYNR 384

RESULT 11
VS12_REV0V STANDARD; PRT; 418 AA.
ID VS12_REV0V
AC P32922;
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 01-OCT-1993 (Rel. 27, Last annotation update)
DS Sigma 2 protein (Core protein).
OS Reovirus (Type 2 / strain D5/Jones).
OC Viruses; dsRNA viruses; Reoviridae; Orthoreovirus.
OX NCBI_TaxId=10885;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE:92015462; PubMed:1920614;
RT Dermody T.S., Schiff L.A., Nibert M.L., Coombs K.M., Fields B.N.;
RT "The S2 gene nucleotide sequences of prototype strains of the three
RT reovirus serotypes: characterization of reovirus core protein sigma
RT 2." ;
RL J. Virol. 65:5721-5731(1991).
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; L19775; AAA47248.1; -.
DR PIR; B41306; FOXRJS.
DR InterPro; IPR004317; Sigma_1_2.
DR Pfam; PF03084; Sigma_1_2; 1.
DR ProDom; PD004419; Sigma_1_2; 1.
KM Core protein.
SQ SEQUENCE 418 AA; 47045 MW; E3DA351F55524213 CRC64;

Query Match 39.7%; Score 46; DB 1; Length 418;
Best Local Similarity 61.5%; Pred. No. 13;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 8 LEAWSNDPOYNK 20
|||::|||::
Db 372 LEAWARDDOYNR 384

RESULT 12
```

```

AC PERB_HUMAN STANDARD; PRT; 3060 AA.
AD _PERB_HUMAN
DT 15-SEP-2003 (Rel. 42, Created)
DT 15-SEP-2003 (Rel. 42, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Bullous pemphigoid antigen 1, isoform 7 (Bullous pemphigoid antigen)
DE (BPA) (Hemidesmosomal plaque protein) (Dystonia musculorum protein)
DE (Fragment).
GN BPA1 OR DMH OR DT OR KIAA0728.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OC NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM 7), AND TISSUE SPECIFICITY.
RC TISSUE=Keratinocytes;
RX MEDLINE=21839111; PubMed=11751855;
RX Okumura M., Yamakawa H., Ohara O., Owaribe K.;
RT "Novel alternative splicings of BPA1 (bullous pemphigoid antigen 1)
RT including the domain structure closely related to MACF (microtubule
RT actin cross-linking factor).";
RL J. Biol. Chem. 277:6682-6687(2002).
RN [2]
RP DISEASE.
RC TISSUE=Keratinocytes;
RX MEDLINE=92011493; PubMed=11717441;
RX Sawamura D., Li K., Chu M.-L., Uitto J.;
RT "Human bullous pemphigoid antigen (BPA1). Amino acid sequences
RT deduced from cloned cDNAs predict biologically important peptide
RT segments and protein domains.";
RL J. Biol. Chem. 266:117784-117790(1991).
CC -1- FUNCTION: Cytoskeletal linker protein. Anchors keratin-containing
CC intermediate filaments to the inner plaque of hemidesmosomes. The
CC proteins may self-aggregate to form filaments or a two-dimensional
CC mesh (By similarity).
CC -1- SUBUNIT: Homodimer (By similarity).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=10;
CC Name=7; Synonyms=SB;
CC IsoId=Q8WXK8-1; Sequence=Displayed;
CC Name=1;
CC IsoId=Q03001-1; Sequence=External;
CC Name=2;
CC IsoId=Q03001-2; Sequence=External;
CC Name=3; Synonyms=1e;
CC IsoId=Q03001-3; Sequence=External;
CC Name=4;
CC IsoId=Q03001-4; Sequence=External;
CC Name=5;
CC IsoId=Q03001-5; Sequence=External;
CC Name=6; Synonyms=EX;
CC IsoId=Q94833-2; Sequence=External;
CC Name=8;
CC IsoId=Q03001-6; Sequence=External;
CC Name=9;
CC IsoId=Q94833-3; Sequence=External;
CC Name=10;
CC IsoId=Q94833-1; Sequence=External;
CC Note=No experimental confirmation available;
CC TISSUE SPECIFICITY: Highly expressed in skeletal muscle and
CC cultured keratinocytes.
CC -1- DISEASE: BPA1 is an autoantigen of bullous pemphigoid
CC (EIM:600088), an autoimmune subepithelial skin blistering disease.
CC -1- SIMILARITY: Belongs to the plakin or cytolinker family.
CC -1- SIMILARITY: Contains 4 spectrin repeats.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way

```


modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@sb-sib.ch).

EMBL, AF400227; AL62062.1; -

MIM, 113810; -

MIM, 600088; -

GO: GO:0005200; F: structural constituent of cytoskeleton; ISS.

GO: GO:0045104; P: intermediate filament cytoskeleton organiza. .; ISS.

InterPro: IPR001101; Plectin_repeat.

InterPro: IPR001452; SH3.

InterPro: IPR002017; Spectrin.

Pfam: PF00681; Plectrin_4.

Pfam: PF00435; Spectrin; 3.

Pfam: PF00018; SH3; 1.

SMART: SM00250; PLEC; 8.

SMART: SM00150; SPC; 4.

SMART: SM00150; SPC; 4.

Antigen; Actin-binding; Coiled coil. Repeat; SH3 domain; Structural protein; Cytoskeleton; Cell adhesion; Calcium; Calcium-binding; Alternative splicing.

DOMAIN 126 160 COILED COIL (POTENTIAL).

DOMAIN 312 447 COILED COIL (POTENTIAL).

DOMAIN 511 546 COILED COIL (POTENTIAL).

DOMAIN 564 616 SH3.

DOMAIN 706 813 COILED COIL (POTENTIAL).

DOMAIN 871 926 COILED COIL (POTENTIAL).

DOMAIN 1013 1042 COILED COIL (POTENTIAL).

DOMAIN 1087 1139 COILED COIL (POTENTIAL).

DOMAIN 1179 1209 COILED COIL (POTENTIAL).

REPEAT 264 341 SPECTRIN 1.

REPEAT 349 444 SPECTRIN 2.

REPEAT 450 550 SPECTRIN 3.

REPEAT 935 1031 SPECTRIN 4.

REPEAT 1208 1257 PLECTIN 1.

REPEAT 1258 1295 PLECTIN 2.

REPEAT 1333 1370 PLECTIN 3.

REPEAT 1401 1446 PLECTIN 4.

REPEAT 1447 1484 PLECTIN 5.

REPEAT 1487 1522 PLECTIN 6.

REPEAT 1523 1560 PLECTIN 7.

REPEAT 1562 1598 PLECTIN 8.

NON_TER 3060 3060

SEQUENCE 3060 AA; 344985 MW; EF68DF6D72C14E2 CRC64;

Query Match 39.7%; Score 46; DB 1; Length 3060;

Best Local Similarity 47.1%; Pred. No. 1.4e+02;

Matches 8; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

3 NHL5LEAMSDNDPQYN 19

2475 SHLSLIASTVDKDPQGN 2491

RESULT 13

PMG2_ARATH STANDARD; PRT; 560 AA.

OC 09M9K1; 16-OCT-2001 (Rel. 40, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Probable 2,3-bisphosphoglycerate-independent phosphoglycerate mutase 2 (EC 5.4.2.1) (Phosphoglyceromutase 2) (BPG-independent PGAM 2) (PGAM-1 2).

GN AT3G08590 OR PT7014.6.

OS Arabidopsis thaliana (Mouse-ear cress).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsids.

OC NCBI_Taxid=3702;

OK [1]

RP SEQUENCE FROM N.A.

RC STRAIN=cv. Columbia;

FX MEDLINE=21016720; PubMed=1130713;

RA Salanoubat M., Lemcke K., Rieger M., Ansoerg M., Unseid M.,
RA Fartmann B., Valle G., Bloeker H., Perez-Alonso M., Obermaier B.,
RA Delensy M., Boutry M., Grivell L.A., Maché R., Puigdomenech P.,
RA De Simone V., Choiane N., Artiguenave F., Robert C., Brotter P.,
RA Wincker P., Catolico L., Weissbach J., Sarin W., Queller F.,
RA Schaefer M., Mueller-Auer S., Gabel C., Fuchs M., Benes V.,
RA Wunbach E., Drzono H., Erle H., Jordan N., Bangert S.,
RA Wiedemann R., Kranz H., Voss H., Holland R., Brandt P., Nakamura G.,
RA Vezzi A., D'Angelo M., Fallavichini A., Toppi S., Salanoubat B.,
RA Conrad A., Hornischer K., Kauer G., Loehner T.H., Nordstark G.,
RA Reichelt J., Scharte M., Schoen O., Bargues M., Terol J., Clément J.,
RA Navarro P., Collado C., Perez-Perez A., Oltmanns B., Duchemin D.,
RA Cooke R., Landie M., Berger-Llano C., Punnett E., Maury D.,
RA de Haan M., Maarse A.C., Alcaraz J.-P., Corté A., Casacuberta E.,
RA Monfort A., Argilou A., Florez M., Liguori R., Vitale D.,
RA Mannhaupt G., Haase D., Schopf H., Rüd S., Zaccaria P., Mewes H.-W.,
RA Mayer K.F.X., Kaul S., Town C.D., Koo H.L., Tallon L.J., Jenkins J.,
RA Rooney T., Rizzo M., Malt A., Utterback T., Fujii C.Y., Shea T.P.,
RA Creasy T.H., Haas B., Malt R., Wu D., Peterson J., Van Aken S.,
RA Pai G., Miltner J., Sellers P., Gill J.E., Feidlyum T.V.,
RA Preuss D., Lin X., Nierman W.C., Salzberg S.L., White O., Venter J.C.,
RA Fraser C.M., Kaneo T., Nakamura Y., Sato S., Kato T., Asamizu E.,
RA Sasamoto S., Kimura T., Ideasa K., Kawashima K., Kishida Y.,
RA Kiyokawa C., Kohara M., Matsumoto M., Matsuno A., Muraki A.,
RA Nakayama S., Nakazaki N., Shino S., Takeuchi C., Wada T.,
RA Matanabe A., Yamada M., Yasuda M., Tabata S.,
RT "Sequence and analysis of chromosome 3 of the plant Arabidopsis
thaliana." ;
RL Nature 408:820-822(2000).

CC -1- FUNCTION: Catalyzes the interconversion of 2-phosphoglycerate and
CC 3-phosphoglycerate (By similarity).

CC -1- CATALYTIC ACTIVITY: 2-phospho-D-glycerate = 3-phospho-D-glycerate.
CC -1- COFACTOR: Binds 2 manganese ions (By similarity).

CC -1- PATHWAY: Glycolysis.

CC -1- SUBUNIT: Monomer (By similarity).

CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).

CC -1- SIMILARITY: BELONGS TO THE BPG-INDEPENDENT PHOSPHOGLYCERATE MUTASE
FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sb-sib.ch).

CC EMBL, AC012562; AAC51361.1; -

DR InterPro: IPR006124; Metalloenzyme.

DR InterPro: IPR005995; Pgm bpd ind.

DR Pfam, PF01676; Metalloenzyme_1.

DR Pfam, PF004429; Pgm bpd ind. 1.

DR TIGRFAMs: TIGR01307; Pgm bpd ind. 1.

KW Isoenzyme; Glycolysis; Metal-binding; Manganese.
KW PHOSPHOGLYCERATE INTERMEDIATE
KW ACT_SITE

FT 82 82

FT METAL 29 29 MANGANESE 2 (BY SIMILARITY).

FT METAL 82 82 MANGANESE 2 (BY SIMILARITY).

FT METAL 431 431 MANGANESE 1 (BY SIMILARITY).

FT METAL 435 435 MANGANESE 1 (BY SIMILARITY).

FT METAL 472 472 MANGANESE 2 (BY SIMILARITY).

FT METAL 473 473 MANGANESE 1 (BY SIMILARITY).

FT METAL 502 502 MANGANESE 2 (BY SIMILARITY).

SEQUENCE 560 AA; 60763 MW; 4D5AC0807AD7B8B CRC64;

Query Match 39.2%; Score 45.5; DB 1; Length 560;

Best Local Similarity 50.0%; Pred. No. 23;

Matches 8; Conservative 4; Mismatches 3; Indels 1; Gaps 1;

5 LSLLEAMSDNDP-QYN 19

25 LVLDDGWSBSPDQYN 40

RESULT 14
 SYM_CLOAB STANDARD: PRT: 644 AA.
 AC 097EM5;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Methionyl-tRNA synthetase (EC 6.1.1.10) (Methionine--tRNA ligase) (MetRS).
 GN METG OR CAC2991.
 OS Clostridium acetobutylicum.
 OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
 OC Clostridium.
 NC NCBI_TaxID=1488;
 RX [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 824 / DSM 792 / VKM B-1787;
 RX MEDLINE=21359325; PubMed=1466286;
 RA Noelling J., Breton G., Omeichenko M.V., Makarova K.S., Zeng Q.,
 RA Gibson R., Lee H.M., Dubois J., Qiu D., Hiltl J., Wolf Y.I.,
 RA Tatsurov R.L., Sabathe P., Doucette-Stamm L., Soucaille P., Daly M.J.,
 RA Bennett G.N., Koonin E.V., Smith D.R.;
 RT "Genome sequence and comparative analysis of the solvent-producing
 bacterium Clostridium acetobutylicum";
 RL J. Bacteriol. 183:4823-4838(2001).
 CC -1- FUNCTION: Is required not only for elongation of protein synthesis
 but also for the initiation of all mRNA translation through
 initiator tRNA (fMet) aminoacylation (By similarity).
 CC -1- CATALYTIC ACTIVITY: ATP + L-methionine + tRNA(Met) = AMP +
 diphosphate + L-methionyl-tRNA(Met).
 CC -1- SUBUNIT: Homodimer (By similarity).
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -1- SIMILARITY: Belongs to class-I aminoacyl-tRNA synthetase family.
 CC MetG subfamily 2A.
 CC -1- SIMILARITY: Contains 1 tRNA-binding domain.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation-
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC
 CC EMBL: AE007796; AAK80932.1; -
 CC PIR: A97268; A97268.
 DR HAMAP: MF_01228; fused. 1.
 DR InterPro: IPR004495; MetG_Cterm.
 DR InterPro: IPR002300; tRNA-synt_1a.
 DR InterPro: IPR001412; tRNA-synt_1.
 DR InterPro: IPR002304; tRNA-synt_mec.
 DR InterPro: IPR002547; tRNA_bind.
 DR Pfam: PF00133; tRNA-synt_1.1.
 DR Pfam: PF01588; tRNA_bind_1.
 DR PIRSF: PIRSF001528; MetRS_dimerising; 1.
 DR PRINTS: PR01041; TRNAYSYNTHET.
 DR TIGRFAMs: TIGR00398; metG; 1.
 DR TIGRFAMs: TIGR00399; metG_C term; 1.
 DR PROSITE: PS00178; AA_TRNA_LIGASE_I; 1.
 DR PROSITE: PSS0886; TRBD; 1.
 DR AMOSyC: tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;
 KW RNA-binding; tRNA-binding; Metal-binding; Zinc; Complete proteome.
 FT SITE 14 24
 FT SITE 299 303
 FT DOMAIN 542 644
 FT METAL 129 129 TRNA-BINDING.
 FT METAL 132 132 TRNA-BINDING.
 FT METAL 146 146 ZINC (BY SIMILARITY).
 FT METAL 149 149 ZINC (BY SIMILARITY).
 FT BINDING 302 302 ZINC (BY SIMILARITY).
 FT BINDING 302 302 ATP (BY SIMILARITY).
 SQ SEQUENCE 644 AA; 73618 MW; 30035F8B94B60A7A CRC64;

Query Match 38.8%; Score 45; DB 1; Length 644;
 Best Local Similarity 50.0%; Pred. No. 32;
 Matches 9; Conservative 3; Mismatches 6; Indels 0; Gaps 0;
 QY 3 NHLSLEAMSNDPQYNK 20
 Db 234 NYITALGYGSDNDELXNK 251
 RESULT 15
 ID CSPI CORGL STANDARD: PRT: 657 AA.
 AC 001377;
 DT 01-JUL-1993 (Rel. 26, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE PSI protein precursor.
 GN CSPI OR CGL2875.
 OS Corynebacterium glutamicum (Brevibacterium flavum).
 OC Bacteria; Actinobacteria; Actinobacteriales; Actinomycetales;
 OC Corynebacteriaceae; Corynebacteriaceae; Corynebacterium.
 NC NCBI_TaxID=1718;
 RX [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 17965 / Melassecola;
 RX MEDLINE=93023863; PubMed=1406274;
 RA Joliffe G., Mathieu L., Hahn V., Bayan N., Duchiron F., Renaud M.,
 RA Sheehy E., Leblon G.;
 RT "Cloning and nucleotide sequence of the cspl gene encoding PSI, one
 of the two major secreted proteins of Corynebacterium glutamicum: the
 deduced N-terminal region of PSI is similar to the Mycobacterium
 antigen 85 complex";
 RL Mol. Microbiol. 6:2349-2362(1992).
 CC [2]
 CC SEQUENCE FROM N.A.
 CC STRAIN=ATCC 13032 / DSM 20300 / NCIB 10025;
 CC Nakagawa S.;
 CC "Complete genomic sequence of Corynebacterium glutamicum ATCC 13032";
 CC Submitted (May-2002) to the EMBL/Genbank/DBJ databases.
 CC -1- FUNCTION: ONE OF THE TWO MAJOR SECRETED PROTEINS.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- SIMILARITY: THE N-TERMINAL IS VERY SIMILAR TO THE COMPLETE
 SEQUENCES OF THE MYCOBACTERIAL ANTIGENS 85-A, 85-B & 85-C.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation-
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC
 CC EMBL: X66078; CAA46877.1; -
 CC EMBL: AF005283; BAC00269.1; -
 CC PIR: S25184; S25184.
 CC HSSP: P31953; IDQ2.
 CC InterPro: IPR000801; Esterase_put.
 CC InterPro: IPR000379; Ser_estrs_site.
 CC Pfam: PF00756; Esterase; 1.
 KW Signal; Complete proteome; Pyrrolidone carboxylic acid.
 FT SIGNAL 1 43
 FT CHAIN 44 657
 FT MOD RES 44 44
 FT CONFLICT 54 54 I -> V (IN REF. 1).
 FT CONFLICT 177 177 V -> I (IN REF. 1).
 FT CONFLICT 189 189 G -> E (IN REF. 1).
 FT CONFLICT 320 320 G -> D (IN REF. 1).
 FT CONFLICT 330 330 A -> T (IN REF. 1).
 FT CONFLICT 350 350 S -> N (IN REF. 1).
 FT CONFLICT 461 461 A -> D (IN REF. 1).
 FT CONFLICT 562 562 E -> D (IN REF. 1).
 FT CONFLICT 596 596 A -> V (IN REF. 1).
 FT CONFLICT 596 596

SQ SEQUENCE 657 AA; 70674 MW; E2C3F1B57A7938A3 CRC64;
 Query Match 38.8%; Score 45; DB 1; Length 657;
 Best Local Similarity 53.8%; Pred. No. 33;
 Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
 QY 9 EAWSDNDPQYNKD 21
 | : | : |
 | : | : |
 Db 286 ERWQENDPKSNVD 298

Search completed: November 13, 2003, 09:45:30
 Job time : 8.56398 secs


```
054178
ID 054178 PRELIMINARY; PRT; 1577 AA.
AC 054178; 054247;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DE 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
DE Glucosyltransferase.
GN GTFg.
OS Streptococcus gordonii Challis.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=29390;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CHALLIS;
RX MEDLINE=96157084; PubMed=8586195;
RA Vickerman M.M., Sulavik M.C., Clewell D.B.;
RT "Molecular analysis of Streptococcus gordonii glucosyltransferase
RT phase variants.";
RL Dev. Biol. Stand. 85:309-314(1995).
RN [2]
RP SEQUENCE OF 1-96 FROM N.A.
RC STRAIN=CHALLIS;
RX MEDLINE=92276337; PubMed=1534326;
RA Sulavik M.C., Tardif G., Clewell D.B.;
RT "Identification of a gene, t9g, which regulates expression of
RT glucosyltransferase and influences the spp phenotype of Streptococcus
RT gordonii Challis.";
RL J. Bacteriol. 174:3577-3586(1992).
DR EMBL; U12643; AAC34483.1; -.
DR EMBL; M89776; AAC26969.1; -.
DR InterPro; IPR002479; CW binding.
DR InterPro; IPR003318; Glyco_hydro_70.
DR Pfam; PF01473; CW binding_1; 18.
DR Pfam; PF02324; Glyco_hydro_70; 1.
KW Transferrase.
SQ SEQUENCE 1577 AA; 177805 MW; 5AE0328DC5E08D18 CRC64;

Query Match 88.8%; Score 103; DB 2; Length 1577;
Best Local Similarity 85.7%; Pred. No. 2.5e-07;
Matches 18; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 AINHLSTLEAWSNDNDPOYNKD 21
DB 548 ALKHLSTLEAWSNDNDPOYNKD 568

RESULT 3
Q55265 PRELIMINARY; PRT; 1577 AA.
ID 055265
AC 055265;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DE 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
DE Glucosyltransferase precursor.
GN GTFM.
OS Streptococcus salivarius.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1304;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=95122197; PubMed=7822030;
RX Simpson C.L., Giffard P.M., Jacques N.A.;
RT "Streptococcus salivarius ATCC 25975 possesses at least two genes
RT coding for primer-independent glucosyltransferases.";
RL Infect. Immun. 63:609-621(1995).
DR EMBL; L35928; AAC41413.1; -.
DR InterPro; IPR004829; Casurface_antigen.
DR InterPro; IPR002479; CW binding.
DR InterPro; IPR003318; Glyco_hydro_70.
DR Pfam; PF01473; CW binding_1; 11.
DR Pfam; PF02324; Glyco_hydro_70; 1.
```

```
DR ProDom; PD153432; Casurface_antigen; 1.
KW Signal; Transferase. 38
FT SIGNAL 1 POTENTIAL.
FT CHAIN 39 1577 GLUCOSYLTRANSFERASE.
SQ SEQUENCE 1577 AA; 175290 MW; 3EFB898A7D3A7BF3 CRC64;

Query Match 77.6%; Score 90; DB 2; Length 1577;
Best Local Similarity 85.7%; Pred. No. 3e-05;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AINHLSTLEAWSNDNDPOYNKD 21
DB 588 AINHLSTLEAWSNDNDPOYNKD 608

RESULT 4
068542
ID 068542 PRELIMINARY; PRT; 1449 AA.
AC 068542;
DT 01-AUG-1998 (TREMBlrel. 07, Created)
DT 01-AUG-1998 (TREMBlrel. 07, Last sequence update)
DE 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
DE Glucosyltransferase N (Fragment).
GN GTFN.
OS Streptococcus salivarius.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1304;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=V1477;
RA Taffe R.I.;
RT "Streptococcus salivarius V1477 gtfN";
RT Submitted (FEB-1998) to the EMBL/Genbank/DBJ databases.
DR EMBL; AF049609; AAC05156.1; -.
DR InterPro; IPR002479; CW binding.
DR InterPro; IPR003318; Glyco_hydro_70.
DR Pfam; PF01473; CW binding_1; 8.
DR Pfam; PF02324; Glyco_hydro_70; 1.
KW Transferrase.
SQ SEQUENCE 1449 AA; 159895 MW; 0700F6D748471BF8 CRC64;

Query Match 73.3%; Score 85; DB 2; Length 1449;
Best Local Similarity 76.2%; Pred. No. 0.00017;
Matches 16; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 AINHLSTLEAWSNDNDPOYNKD 21
DB 536 AINHLSTLEAWSNDNDPOYNKD 556

RESULT 5
Q55264 PRELIMINARY; PRT; 1449 AA.
ID 055264
AC 055264;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DE 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
DE Glucosyltransferase precursor.
GN GTFL.
OS Streptococcus salivarius.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1304;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=95122197; PubMed=7822030;
RX Simpson C.L., Giffard P.M., Jacques N.A.;
RT "Streptococcus salivarius ATCC 25975 possesses at least two genes
RT coding for primer-independent glucosyltransferases.";
RL Infect. Immun. 63:609-621(1995).
DR EMBL; L35495; AAC41412.1; -.
```

DR InterPro; IPR002479; CW binding.
 DR InterPro; IPR003318; Glyco_hydro_70.
 DR Pfam; PF01473; CW_binding_1; 8.
 DR Pfam; PF02324; Glyco_hydro_70; 1.
 KW Signal; Transferase.
 FT SIGNAL 1 35 POTENTIAL.
 FT CHAIN 36 1449 GLYCOSYLTRANSFERASE.
 SQ SEQUENCE 1449 AA; 159984 MW; DD62F07306E86A46 CRC64;

Query Match 73.3%; Score 85; DB 2; Length 1449;
 Best Local Similarity 76.2%; Pred. No. 0.00017;
 Matches 16; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 AINHLSTLEAMSDNDPOYKND 21
 DB 536 ALKHLSTLEAMSHNDAYTNE 556

RESULT 6
 Q9WKJ4 PRELIMINARY; PRT; 1338 AA.

AC Q9WKJ4; 01-NOV-1999 (TREMBlrel. 12, Created)
 DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
 DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
 DE GTP-S.
 GN Streptococcus criceti.
 OS Streptococcus criceti.
 OC Placmid pAM1.
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 OX NCBI_TaxID=1333;
 RN [1] SEQUENCE FROM N.A.
 RC STRAIN=H-6;
 RA Inoue M., Fukui K., Miyagi A.;
 RT "Subcloned glucosyltransferase (gfts and gftT) genes";
 RL Submitted (Mar-1999) to the EMBL/Genbank/DBJ databases.
 DR EMBL; AB028123; BAA77236.1;
 DR InterPro; IPR002479; CW_binding.
 DR InterPro; IPR003318; Glyco_hydro_70.
 DR Pfam; PF01473; CW_binding_1; 10.
 DR Pfam; PF02324; Glyco_hydro_70; 1.
 KW Placmid.
 SQ SEQUENCE 1338 AA; 148558 MW; 0A90C6E10E15D99B CRC64;

Query Match 69.8%; Score 81; DB 2; Length 1338;
 Best Local Similarity 76.2%; Pred. No. 0.00066;
 Matches 16; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 AINHLSTLEAMSDNDPOYKND 21
 DB 437 AIDHLSTLEAMSGNDNDYKND 457

RESULT 7

Q8KZL5 PRELIMINARY; PRT; 1554 AA.

AC Q8KZL5; 01-OCT-2002 (TREMBlrel. 22, Created)
 DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)
 DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
 DE Glucosyltransferase.
 GN GTFU.
 OS Streptococcus sobrinus.
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 OX NCBI_TaxID=1310;
 RN [1] SEQUENCE FROM N.A.
 RP MEDLINE=21958684; PubMed=11960691;
 RA Hanada N., Fukushima K., Nomura Y., Sanpuku H., Hayakawa M.,
 Mukasa H., Shiroza T., Abiko Y.;

RT "Cloning and nucleotide sequence analysis of the Streptococcus
 RT sobrinus gtfu gene that produces a highly branched water-soluble
 RT glucan";
 RL Biochim. Biophys. Acta 1570:75-79 (2002).

DR EMBL; AB089438; BAC07265.1;
 DR InterPro; IPR002479; CW_binding.
 DR InterPro; IPR003318; Glyco_hydro_70.
 DR Pfam; PF01473; CW_binding_1; 6.
 DR Pfam; PF02324; Glyco_hydro_70; 1.
 KW Transferase.
 SQ SEQUENCE 1554 AA; 171676 MW; 6981BC1DAE24A73 CRC64;

Query Match 69.8%; Score 81; DB 2; Length 1554;
 Best Local Similarity 71.4%; Pred. No. 0.00078;
 Matches 15; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 AINHLSTLEAMSDNDPOYKND 21
 DB 485 ALAHLSTLEAMSLNDNOYKND 505

RESULT 8

Q9L466 PRELIMINARY; PRT; 1477 AA.

AC Q9L466; 01-OCT-2000 (TREMBlrel. 15, Created)
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
 DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
 DE Dextranucrase (EC 2.4.1.5).
 GN DSRG.
 OS Leuconostoc mesenteroides.
 OC Bacteria; Firmicutes; Lactobacillales; Leuconostoc.
 OX NCBI_TaxID=1245;
 RN [1] SEQUENCE FROM N.A.
 RC STRAIN=NRL B-1355;
 RA Arguello-Morales M.A., Remaud-Simeon M., Pizut S., Sarcabal P.,
 RA Willemot R.M., Monsan P.;
 RT "Sequence analysis of the gene encoding alternanucrase, a sucrose
 RT glucosyltransferase from Leuconostoc mesenteroides NRL B-1355";
 RL Submitted (Oct-1999) to the EMBL/Genbank/DBJ databases.
 DR EMBL; AJ250172; CAB76565.1;
 DR InterPro; IPR002479; CW_binding.
 DR InterPro; IPR003318; Glyco_hydro_70.
 DR Pfam; PF01473; CW_binding_1; 14.
 DR Pfam; PF02324; Glyco_hydro_70; 1.
 KW Glycosyltransferase; Transferase.
 SQ SEQUENCE 1477 AA; 164887 MW; B6F5710DEDFC831 CRC64;

Query Match 67.2%; Score 78; DB 2; Length 1477;
 Best Local Similarity 71.4%; Pred. No. 0.0022;
 Matches 15; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 AINHLSTLEAMSDNDPOYKND 21
 DB 532 ANQHLSTLEAMSHNDPEYKND 552

RESULT 9

Q9EZHS PRELIMINARY; PRT; 1508 AA.

AC Q9EZHS; 01-MAR-2001 (TREMBlrel. 16, Created)
 DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
 DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
 DE Dextranucrase Derb742.
 GN DSRB742.
 OS Leuconostoc mesenteroides.
 OC Bacteria; Firmicutes; Lactobacillales; Leuconostoc.
 OX NCBI_TaxID=1245;
 RN [1] SEQUENCE FROM N.A.
 RP STRAIN=B-742CB;

RA Kim H.-S., Kim D., Ryu H.-J., Roby J.F.
 RT "Leucostoc mesenteroides B-742CB, a dextranucrase gene."
 RL Submitted (Aug-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: A2294469; ANG38021.1; -
 DR InterPro: IPR002479; CW_binding_70.
 DR InterPro: IPR003318; Glyco_hydro_70.
 DR Pfam: PF01473; CW_binding_1; 14.
 DR Pfam: PF02324; Glyco_hydro_70; 1.
 SQ SEQUENCE 1508 AA; 168542 MW; E2FCFA0F87AE4F3A CRC64;

Query Match 67.2%; Score 78; DB 2; Length 1508;
 Best Local Similarity 71.4%; Pred. No. 0.0023;
 Matches 15; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 AINHLSTLEAWSNDNDPOYNKD 21
 DB 563 ANOHLSTLEAWSNDNDPEYVKD 583

RESULT 10

052224

ID 052224; PRELIMINARY; PRT; 1508 AA.

AC 052224; (TREMBlrel. 06, Created)
 DT 01-JUN-1998 (TREMBlrel. 06, Last sequence update)
 DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
 DE Glucosyltransferase (EC 2.4.1.5).
 GN DSRB.
 OS Leucostoc mesenteroides.
 OC Bacteria; Firmicutes; Lactobacillales; Leucostoc.
 OX NCBI_TaxID=1245;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=NRRL B-1299;
 RA Monchois V., Renaud-Simeon M., Monsan P., Willemot R.M.;
 RT "Cloning and sequencing of a gene coding for an extracellular
 dextranucrase (DSRB) from Leucostoc mesenteroides NRRL B-1299
 RT extralisting only a a(1-6) glucan."
 RL FEMS Microbiol. Lett. 0:0-0(1998).
 DR EMBL: AF030129; AAB95453.1; -
 DR InterPro: IPR002479; CW_binding_70.
 DR InterPro: IPR003318; Glyco_hydro_70.
 DR Pfam: PF01473; CW_binding_1; 14.
 DR Pfam: PF02324; Glyco_hydro_70; 1.
 KM Glycosyltransferase; Transferase.
 SQ SEQUENCE 1508 AA; 168511 MW; E70CEGB57A70D1F0 CRC64;

Query Match 67.2%; Score 78; DB 2; Length 1508;
 Best Local Similarity 71.4%; Pred. No. 0.0023;
 Matches 15; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 AINHLSTLEAWSNDNDPOYNKD 21
 DB 563 ANOHLSTLEAWSNDNDPEYVKD 583

RESULT 11

09WXJ5

ID 09WXJ5; PRELIMINARY; PRT; 1512 AA.

AC 09WXJ5; (TREMBlrel. 12, Created)
 DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
 DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
 DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
 DE GTF-S.
 GN GTF.
 OS Streptococcus criceti.
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 OX NCBI_TaxID=1333;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=HS-6;

RA Inoue M., Fukui K., Miyagi A.
 RT "S.cricetus glucosyltransferase (gtfs and gtfI) genes."
 RL Submitted (Mar-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AB026123; BAA77237.1; -
 DR InterPro: IPR002479; CW_binding_70.
 DR InterPro: IPR003318; Glyco_hydro_70.
 DR Pfam: PF01473; CW_binding_1; 14.
 DR Pfam: PF02324; Glyco_hydro_70; 1.
 KM PfamId.
 SQ SEQUENCE 1512 AA; 167145 MW; 4C03D9C8C601FC14 CRC64;

Query Match 65.5%; Score 76; DB 2; Length 1512;
 Best Local Similarity 66.7%; Pred. No. 0.0047;
 Matches 14; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 AINHLSTLEAWSNDNDPOYNKD 21
 DB 485 ALAHLSTLEAWSNDNYNOD 505

RESULT 12

059983

ID 059983; PRELIMINARY; PRT; 1590 AA.

AC 059983;
 DT 01-NOV-1996 (TREMBlrel. 01, Created)
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
 DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
 DE Glucosyltransferase-I precursor (EC 2.4.1.5).
 GN GTFI.
 OS Streptococcus sobrinus.
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OX NCBI_TaxID=1310;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=QM2176;
 RA MEDLINE-94146405; PubMed-8312602;
 RA Sato S., Inoue M., Handa N., Aizawa Y., Isobe Y., Katayama T.;
 RT "DNA sequence of the glucosyltransferase gene of serotype d
 RT Streptococcus sobrinus."
 RL DNA Seq. 4:19-27(1993).
 DR EMBL: D13858; BAA02976.1; -
 DR InterPro: IPR002479; CW_binding_70.
 DR InterPro: IPR003318; Glyco_hydro_70.
 DR Pfam: PF01473; CW_binding_1; 16.
 DR Pfam: PF02324; Glyco_hydro_70; 1.
 KM Glycosyltransferase; Signal; Transferase.
 FT SIGNAL.
 FT CHAIN 1 38
 FT SIGNAL 39 1590
 FT CHAIN 1 38
 SQ SEQUENCE 1590 AA; 175955 MW; C3C83A57CF3C2B0E CRC64;

Query Match 63.8%; Score 74; DB 2; Length 1590;
 Best Local Similarity 66.7%; Pred. No. 0.01;
 Matches 14; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 1 AINHLSTLEAWSNDNDPOYNKD 21
 DB 477 ANHWSTLEAWSNDNDPTYLHD 497

RESULT 13

055263

ID 055263; PRELIMINARY; PRT; 1590 AA.

AC 055263;
 DT 01-NOV-1996 (TREMBlrel. 01, Created)
 DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)
 DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
 DE GTF-I.
 GN GLUCOSYLTRANSFERASE.
 OS Streptococcus sobrinus.
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 OX NCBI_TaxID=1310;

```

RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 33478;
RA Sato S.;
RT "DNA and amino-acid sequences of water-insoluble-glucan synthetase
   produced from Streptococcus sobrinus ATCC 33478."
RL Ann. Kagoshima Univ. Dental School 16:23-29(1996).
DR EMBL; D63570; BAA09792.1; -.
DR InterPro; IPR002479; CW binding.
DR InterPro; IPR003318; Glyco_hydro_70.
DR Pfam; PF01473; CW binding_1; 15.
DR Pfam; PF02324; Glyco_hydro_70; 1.
SQ SEQUENCE 1590 AA; 176058 MW; 9DF7A3F2C6E4FD43 CRC64;

Query Match
Best Local Similarity 63.8%; Score 74; DB 2; Length 1590;
Matches 14; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 1 AINHLSTLEAWSNDNDPQYKND 21
DB 477 ANNHVSIVEAWSNDNDPTPLHD 497

RESULT 14
Q9LCU7 PRELIMINARY; PRT; 1016 AA.
ID Q9LCU7;
AC Q9LCU7;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-OCT-2002 (TREMBLrel. 22, Last annotation update)
DE Dextranucrase.
GN DSRT.
OS Leuconostoc mesenteroides.
OC Bacteria; Firmicutes; Lactobacillales; Leuconostoc.
OX NCBI_TaxID=1245;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NRRL B-512F;
RX MEDLINE=20169623; PubMed=10705445;
RA Funane K., Mizuno K., Takahara H., Kobayashi M.;
RT "Gene encoding a dextranucrase-like protein in Leuconostoc
   mesenteroides NRRL B-512F."
RL BioSci. Biotechnol. Biochem. 64:29-38(2000).
DR EMBL; AB020020; BAA90527.1; -.
DR HSSP; P06278; 1VJ5.
DR InterPro; IPR003318; Glyco_hydro_70.
DR Pfam; PF02324; Glyco_hydro_70; 1.
SQ SEQUENCE 1016 AA; 110344 MW; 88966FDE13CCCB47 CRC64;

Query Match
Best Local Similarity 62.9%; Score 73; DB 2; Length 1016;
Matches 13; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 4 HUSILFEAWSNDNDPQYKND 21
DB 551 HVSILFEAWSNDNDPEYKND 568

RESULT 15
Q00600 PRELIMINARY; PRT; 1518 AA.
AC Q00600;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-OCT-2002 (TREMBLrel. 22, Last annotation update)
DE Glucosyltransferase I (EC 2.4.1.5) (GTF) (Dextranucrase) (Sucrose 6-
   glucosyltransferase).
GN GTFV.
OS Streptococcus salivarius.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1304;
RN [1]

```

```

RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 25975; PubMed=1838391;
RX MEDLINE=92148377;
RA Giffard P.M., Simpson C.L., Milward C.P., Jacques N.A.;
RT "Molecular characterization of a cluster of at least two-
   glucosyltransferase genes in Streptococcus salivarius ATCC 25975."
RL J. Gen. Microbiol. 137:2577-2593(1991).
CC CC
CC FUNCTION: PRODUCTION OF EXTRACELLULAR GLUCANS, THAT ARE THOUGHT TO
CC PLAY A KEY ROLE IN THE DEVELOPMENT OF THE DENTAL PLAQUE BECAUSE OF
CC THEIR ABILITY TO ADHERE TO SMOOTH SURFACES AND MEDIANE THE
CC AGGREGATION OF BACTERIAL CELLS AND FOOD DEBRIS.
CC -1- CATALYTIC ACTIVITY: SUCROSE + ((1-6)-ALPHA-D-GLUCOSYL) (N) = D-
CC FRUCTOSE + ((1-6)-ALPHA-D-GLUCOSYL) (N+1).
CC -1- SUBCELLULAR LOCATION: EXTRACELLULAR.
CC -1- DISEASE: DENTAL CARIES.
DR EMBL; Z11873; CAAT7900.1; -.
DR EMBL; M64111; AA26896.1; -.
DR InterPro; IPR002479; CW binding.
DR InterPro; IPR003318; Glyco_hydro_70.
DR Pfam; PF01473; CW binding_1; 13.
DR Pfam; PF02324; Glyco_hydro_70; 1.
KW Transferase; Glucosyltransferase; Repeat; Dental caries.
FT DOMAIN 1307 1482 6 DIRECT REPEATS.
FT REPEAT 1307 1338 REPEAT 1.
FT REPEAT 1339 1352 REPEAT 2.
FT REPEAT 1372 1403 REPEAT 3.
FT REPEAT 1404 1417 REPEAT 4.
FT REPEAT 1437 1468 REPEAT 5.
FT REPEAT 1469 1482 REPEAT 6.
SQ SEQUENCE 1518 AA; 167730 MW; DAA41F717098B59A CRC64;

Query Match
Best Local Similarity 58.6%; Score 68; DB 2; Length 1518;
Matches 12; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 AINHLSTLEAWSNDNDPQYKND 19
DB 501 ALAHVSIVEAWSNDNDHNHN 519

```

Search completed: November 13, 2003, 09:44:03
 Job time : 33.6493 sec

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Comugen Ltd.

CM protein - protein search, using sw model

Run on: November 13, 2003, 09:31:40 ; Search time 13.6351 Seconds
(without alignments)
65.165 Million cell updates/sec

Title: US-09-290-049a-11

Perfect score: 116

Sequence: 1 AINHLSTLEAWSNDNDPQYNKD 21

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: Issued Patents AA:*

1: /cgn2_6/ptodata/2/1aa/5A.COMB.pep:*

2: /cgn2_6/ptodata/2/1aa/5B.COMB.pep:*

3: /cgn2_6/ptodata/2/1aa/6A.COMB.pep:*

4: /cgn2_6/ptodata/2/1aa/6B.COMB.pep:*

5: /cgn2_6/ptodata/2/1aa/PCTUS.COMB.pep:*

6: /cgn2_6/ptodata/2/1aa/backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	116	100.0	545	4	US-09-604-957-4
2	116	100.0	1430	3	US-09-008-172-2
3	116	100.0	1430	3	US-09-210-361-6
4	116	100.0	1430	4	US-09-740-274-6
5	90	77.6	1577	2	US-08-793-824-2
6	75	64.7	1475	3	US-09-007-999-2
7	75	64.7	1475	3	US-09-210-361-2
8	75	64.7	1475	4	US-09-740-274-2
9	68	58.6	523	4	US-09-604-957-5
10	66	56.9	1375	3	US-09-210-361-4
11	66	56.9	1375	4	US-09-740-274-4
12	64	55.2	584	4	US-09-604-957-6
13	64	55.2	2057	4	US-09-499-203-2
14	53.5	46.1	535	4	US-09-604-957-7
15	53.5	46.1	1278	4	US-09-604-957-3
16	49	42.2	623	4	US-09-205-258-482
17	49	42.2	704	4	US-09-205-258-422
18	47	40.5	521	4	US-09-107-532A-6431
19	46.5	40.1	327	4	US-09-107-532A-6181
20	46	39.7	31	4	US-09-308-345A-42
21	46	39.7	703	3	US-09-422-869-27
22	46	39.7	703	4	US-09-308-345A-49
23	46	39.7	10182	4	US-09-134-001C-3159
24	45	38.8	151	4	US-09-198-452A-815
25	45	38.8	324	3	US-08-508-761B-31
26	45	38.8	657	3	US-08-508-761B-2
27	44.5	38.4	259	1	US-08-277-231A-3

ALIGNMENTS

28	44.5	38.4	259	2	US-08-473-750-6	Sequence 6, Appli
29	44.5	38.4	259	2	US-08-477-326-6	Sequence 6, Appli
30	44	37.9	467	4	US-09-134-001C-4200	Sequence 4200, Ap
31	43.5	37.5	197	3	US-08-897-020-7	Sequence 7, Appli
32	43.5	37.5	197	4	US-09-350-823-7	Sequence 7, Appli
33	43.5	37.5	691	4	US-09-313-942-20	Sequence 20, Appli
34	43.5	37.5	694	4	US-09-313-942-18	Sequence 18, Appli
35	43.5	37.5	694	4	US-09-313-942-22	Sequence 22, Appli
36	43.5	37.5	784	4	US-09-313-942-30	Sequence 30, Appli
37	43.5	37.5	793	4	US-09-313-942-32	Sequence 32, Appli
38	42	36.2	205	2	US-08-684-024-6	Sequence 6, Appli
39	42	36.2	205	3	US-09-145-868-6	Sequence 6, Appli
40	42	36.2	1001	1	US-07-797-556-6	Sequence 6, Appli
41	42	36.2	1001	1	US-07-943-843-2	Sequence 2, Appli
42	42	36.2	1097	1	US-08-347-003-2	Sequence 30, Appli
43	42	36.2	1097	1	US-07-943-843-6	Sequence 6, Appli
44	42	36.2	1097	1	US-08-347-003-6	Sequence 6, Appli
45	41.5	35.8	720	3	US-09-296-284-25	Sequence 25, Appli

RESULT 1
US-09-604-957-4
Sequence 4, Application US/09604957

Patent No. 6486314

GENERAL INFORMATION:

APPLICANT: VAN GEL-SCHUTTEN, GERRITJINA HENDRIKA

APPLICANT: DIJKHUIZEN, LUBBERT

APPLICANT: RAHAKOU, HAKIM

APPLICANT: LEBER, ROBERT-JAN

TITLE OF INVENTION: NOVEL GLUCOSYLTRANSFERASE AND NOVEL GLUCAN

FILE REFERENCE: BO 43388

CURRENT APPLICATION NUMBER: US/09/604,957

CURRENT FILING DATE: 2000-06-28

PRIOR APPLICATION NUMBER: 00201871.1

PRIOR FILING DATE: 2000-05-25

NUMBER OF SEQ ID NOS: 17

SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 4

LENGTH: 545

TYPE: PRT

ORGANISM: Streptococcus mutans

US-09-604-957-4

Query Match

Best Local Similarity 100.0%; Score 116; DB 4; Length 545;

Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AINHLSTLEAWSNDNDPQYNKD 21

Db 75 AINHLSTLEAWSNDNDPQYNKD 95

RESULT 2

US-09-008-172-2

Sequence 2, Application US/09008172

Patent No. 6127602

GENERAL INFORMATION:

APPLICANT: Nichols, Scott E.

TITLE OF INVENTION: Substitutes for Modified Starch and

FILE REFERENCE: 0358D

CURRENT APPLICATION NUMBER: US/09/008,172

CURRENT FILING DATE: 1998-01-16

EARLIER APPLICATION NUMBER: 08/482,711

EARLIER FILING DATE: 1995-06-07

NUMBER OF SEQ ID NOS: 2

SOFTWARE: FastSeq for Windows Version 3.0

SEQ ID NO 2

LENGTH: 1430

TYPE: PRT


```

; ORGANISM: Streptococcus mutans
US-09-008-172-2
Query Match          100.0%; Score 116; DB 3; Length 1430;
Best Local Similarity 100.0%; Pred. No. 1,4e-10;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AINHSTLEAMSDNDPQYKND 21
Db 495 AINHSTLEAMSDNDPQYKND 515

RESULT 3
US-09-210-361-6
; Sequence 6, Application US/09210361
; Patent No. 6284479
; GENERAL INFORMATION:
; APPLICANT: Nichols, Scott E.
; TITLE OF INVENTION: Substitutes for Modified Starches and
; FILE REFERENCE: 0357CR
; CURRENT APPLICATION NUMBER: US/09/210,361
; EARLIER FILING DATE: 1998-12-11
; EARLIER APPLICATION NUMBER: 09/007,999
; EARLIER FILING DATE: 1998-01-16
; EARLIER APPLICATION NUMBER: 08/478,704
; EARLIER FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 09/009,620
; EARLIER FILING DATE: 1998-01-20
; EARLIER APPLICATION NUMBER: 08/485,243
; EARLIER FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 09/008,172
; EARLIER FILING DATE: 1998-01-16
; EARLIER APPLICATION NUMBER: 08/482,711
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 6
; LENGTH: 1430
; TYPE: PRT
; ORGANISM: Streptococcus mutans
US-09-210-361-6

Query Match          100.0%; Score 116; DB 3; Length 1430;
Best Local Similarity 100.0%; Pred. No. 1,4e-10;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AINHSTLEAMSDNDPQYKND 21
Db 495 AINHSTLEAMSDNDPQYKND 515

RESULT 4
US-09-740-274-6
; Sequence 6, Application US/09740274
; Patent No. 6465203
; GENERAL INFORMATION:
; APPLICANT: Nichols, Scott E.
; TITLE OF INVENTION: Glucan-containing Compositions and Paper
; FILE REFERENCE: 0357CRD
; CURRENT APPLICATION NUMBER: US/09/740,274
; EARLIER FILING DATE: 2000-12-19
; EARLIER APPLICATION NUMBER: 09/210,361
; EARLIER FILING DATE: 1998-12-11
; EARLIER APPLICATION NUMBER: 09/007,999
; EARLIER FILING DATE: 1998-01-16
; EARLIER APPLICATION NUMBER: 08/478,704
; EARLIER FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 09/009,620
; EARLIER FILING DATE: 1998-01-20
; EARLIER APPLICATION NUMBER: 08/485,243
; EARLIER FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 09/008,172
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 6
; LENGTH: 1430
; TYPE: PRT
; ORGANISM: Streptococcus mutans
US-09-210-361-6

Query Match          100.0%; Score 116; DB 3; Length 1430;
Best Local Similarity 100.0%; Pred. No. 1,4e-10;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AINHSTLEAMSDNDPQYKND 21
Db 495 AINHSTLEAMSDNDPQYKND 515
```

```

; PRIOR FILING DATE: 1998-01-16
; PRIOR APPLICATION NUMBER: 08/482,711
; PRIOR FILING DATE: 1995-06-07
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 6
; LENGTH: 1430
; TYPE: PRT
; ORGANISM: Streptococcus mutans
US-09-740-274-6

Query Match          100.0%; Score 116; DB 4; Length 1430;
Best Local Similarity 100.0%; Pred. No. 1,4e-10;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AINHSTLEAMSDNDPQYKND 21
Db 495 AINHSTLEAMSDNDPQYKND 515

RESULT 5
US-08-793-824-2
; Sequence 2, Application US/08793824
; Patent No. 5981838
; GENERAL INFORMATION:
; APPLICANT: Simpson, Christine Lynn
; APPLICANT: Giffard, Philip Morrison
; APPLICANT: Jacques, Nicholas Anthony
; TITLE OF INVENTION: Genetic Manipulation of Plants to
; TITLE OF INVENTION: Increase Stored Carbohydrates
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Griffith Hack & Co
; STREET: Level 8, 168 Walker Street
; CITY: No. 5981838ch Sydney
; STATE: New South Wales
; COUNTRY: Australia
; ZIP: 2060
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/793,824
; FILING DATE:
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: AU PM7643
; FILING DATE: 24-AUG-1994
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 61 2 9957 5944
; TELEFAX: 61 2 957 6288
; TELEX: 26547
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1577 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: not relevant
; MOLECULE TYPE: protein
; ORIGINAL SOURCE:
; ORGANISM: Streptococcus salivarius
US-08-793-824-2

Query Match          77.6%; Score 90; DB 2; Length 1577;
Best Local Similarity 85.7%; Pred. No. 3,1e-06;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AINHSTLEAMSDNDPQYKND 21
Db 588 AINHSTLEAMSDNDPQYKND 608
```

```
RESULT 6
US-09-007-999-2
; Sequence 2, Application US/09007999
; Patent No. 6087559
; GENERAL INFORMATION:
; APPLICANT: Nichols, Scott E.
; TITLE OF INVENTION: Substitutes for Modified Starch and
; FILE REFERENCE: 0356D
; CURRENT APPLICATION NUMBER: US/09/007,999
; EARLIER FILING DATE: 1998-01-16
; EARLIER FILING DATE: 1995-06-07
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 1475
; TYPE: PRT
; ORGANISM: Streptococcus mutans
US-09-007-999-2

Query Match
Best Local Similarity 64.7%; Score 75; DB 3; Length 1475;
Matches 15; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 AINHSLIEAWSNDNDPQYKND 21
DB 481 ANDHLSILEAWSNDNDTPYLHD 501

RESULT 7
US-09-210-361-2
; Sequence 2, Application US/09210361
; Patent No. 6284479
; GENERAL INFORMATION:
; APPLICANT: Nichols, Scott E.
; TITLE OF INVENTION: Substitutes for Modified Starches and
; FILE REFERENCE: 0357CR
; CURRENT APPLICATION NUMBER: US/09/210,361
; EARLIER FILING DATE: 1998-12-11
; EARLIER APPLICATION NUMBER: 09/007,999
; EARLIER FILING DATE: 1998-01-16
; EARLIER APPLICATION NUMBER: 08/478,704
; EARLIER FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 09/009,620
; EARLIER FILING DATE: 1998-01-20
; EARLIER APPLICATION NUMBER: 08/485,243
; EARLIER FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 09/008,172
; EARLIER FILING DATE: 1998-01-16
; EARLIER APPLICATION NUMBER: 08/482,711
; EARLIER FILING DATE: 1995-06-07
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 1475
; TYPE: PRT
; ORGANISM: Streptococcus mutans
US-09-210-361-2

Query Match
Best Local Similarity 64.7%; Score 75; DB 3; Length 1475;
Matches 15; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 AINHSLIEAWSNDNDPQYKND 21
DB 481 ANDHLSILEAWSNDNDTPYLHD 501

RESULT 8
US-09-740-274-2
```

```
; Sequence 2, Application US/09740274
; Patent No. 6465203
; GENERAL INFORMATION:
; APPLICANT: Nichols, Scott E.
; TITLE OF INVENTION: Glucan-containing Compositions and Paper
; FILE REFERENCE: 0357CRD
; CURRENT APPLICATION NUMBER: US/09/740,274
; CURRENT FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: 09/210,361
; PRIOR FILING DATE: 1998-12-11
; PRIOR APPLICATION NUMBER: 09/007,999
; PRIOR FILING DATE: 1998-01-16
; PRIOR APPLICATION NUMBER: 08/478,704
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: 09/009,620
; PRIOR FILING DATE: 1998-01-20
; PRIOR APPLICATION NUMBER: 08/485,243
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: 09/008,172
; PRIOR FILING DATE: 1998-01-16
; PRIOR APPLICATION NUMBER: 08/482,711
; PRIOR FILING DATE: 1995-06-07
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 1475
; TYPE: PRT
; ORGANISM: Streptococcus mutans
US-09-740-274-2

Query Match
Best Local Similarity 64.7%; Score 75; DB 4; Length 1475;
Matches 15; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 AINHSLIEAWSNDNDPQYKND 21
DB 481 ANDHLSILEAWSNDNDTPYLHD 501

RESULT 9
US-09-604-957-5
; Sequence 5, Application US/09604957
; Patent No. 6486314
; GENERAL INFORMATION:
; APPLICANT: VAN GEEL-SCHUTTEN, GERRITDINA HENDRIKA
; APPLICANT: DIKHUIZEN, LOBBERT
; APPLICANT: RAHOUTI, HAKIM
; APPLICANT: LEER, ROBERT-JAN
; TITLE OF INVENTION: NOVEL GLUCOSYLTRANSFERASE AND NOVEL GLUCAN
; FILE REFERENCE: BO 43388
; CURRENT APPLICATION NUMBER: US/09/604,957
; CURRENT FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: 00201871.1
; PRIOR FILING DATE: 2000-05-25
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 523
; TYPE: PRT
; ORGANISM: Leuconostoc mesenteroides
US-09-604-957-5

Query Match
Best Local Similarity 58.6%; Score 68; DB 4; Length 523;
Matches 14; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 AINHSLIEAWSNDNDPQYKND 21
DB 75 ANOHLSDLEWSDNDPQYKND 95

RESULT 10
US-09-210-361-4
```

```
; Sequence 4, Application US/09210361
; Patent No. 6284479
; GENERAL INFORMATION:
; APPLICANT: Nichols, Scott E.
; TITLE OF INVENTION: Substitutes for Modified Starches and
; TITLE OF INVENTION: Latexes in Paper Manufacture
; FILE REFERENCE: 0357CR
; CURRENT APPLICATION NUMBER: US/09/210,361
; EARLIER FILING DATE: 1998-12-11
; EARLIER APPLICATION NUMBER: 09/007,999
; EARLIER FILING DATE: 1998-01-16
; EARLIER APPLICATION NUMBER: 08/478,704
; EARLIER FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 09/009,620
; EARLIER FILING DATE: 1998-01-20
; EARLIER APPLICATION NUMBER: 08/485,243
; EARLIER FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 09/008,172
; EARLIER FILING DATE: 1998-01-16
; EARLIER APPLICATION NUMBER: 08/482,711
; EARLIER FILING DATE: 1995-06-07
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 1375
; TYPE: PRT
; ORGANISM: streptococcus mutans
US-09-210-361-4
```

```
Query Match          56.9%; Score 66; DB 3; Length 1375;
Best Local Similarity 66.7%; Pred. No. 0.025;
Matches 14; Conservative 1; Mismatches 6; Indels 0; Gaps 0;
```

```
QY 1 AINHLSTLEAMSDNDPOYND 21
DB 507 ANKHLSTLEAMSYNDTPYLHD 527
```

```
RESULT 11
US-09-740-274-4
; Sequence 4, Application US/09740274
; Patent No. 6465203
; GENERAL INFORMATION:
; APPLICANT: Nichols, Scott E.
; TITLE OF INVENTION: Glucan-containing Compositions and Paper
; FILE REFERENCE: 0357CRD
; CURRENT APPLICATION NUMBER: US/09/740,274
; CURRENT FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: 09/210,361
; PRIOR FILING DATE: 1998-12-11
; PRIOR APPLICATION NUMBER: 09/007,999
; PRIOR FILING DATE: 1998-01-16
; PRIOR APPLICATION NUMBER: 08/478,704
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: 09/009,620
; PRIOR FILING DATE: 1998-01-20
; PRIOR APPLICATION NUMBER: 08/485,243
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: 09/008,172
; PRIOR FILING DATE: 1998-01-16
; PRIOR APPLICATION NUMBER: 08/482,711
; PRIOR FILING DATE: 1995-06-07
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 1375
; TYPE: PRT
; ORGANISM: streptococcus mutans
US-09-740-274-4
```

```
Query Match          56.9%; Score 66; DB 4; Length 1375;
Best Local Similarity 66.7%; Pred. No. 0.025;
Matches 14; Conservative 1; Mismatches 6; Indels 0; Gaps 0;
```

```
QY 1 AINHLSTLEAMSDNDPOYND 21
DB 507 ANKHLSTLEAMSYNDTPYLHD 527
```

```
RESULT 12
US-09-604-957-6
; Sequence 6, Application US/09604957
; Patent No. 6486314
; GENERAL INFORMATION:
; APPLICANT: VAN GEEL-SCHUTTEN, GERRITJINA HENDRIKA
; APPLICANT: DIJKHUIZEN, LOUBERT
; APPLICANT: RAHROUT, HAKIM
; APPLICANT: LEER, ROBERT-JAN
; TITLE OF INVENTION: NOVEL GLUCOSYLTRANSFERASE AND NOVEL GLUCAN
; FILE REFERENCE: BO 43388
; CURRENT APPLICATION NUMBER: US/09/604,957
; CURRENT FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: 00201871.1
; PRIOR FILING DATE: 2000-05-25
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 584
; TYPE: PRT
; ORGANISM: Leuconostoc mesenteroides
US-09-604-957-6
```

```
Query Match          55.2%; Score 64; DB 4; Length 584;
Best Local Similarity 66.7%; Pred. No. 0.019;
Matches 12; Conservative 1; Mismatches 5; Indels 0; Gaps 0;
```

```
QY 1 AINHLSTLEAMSDNDPOY 18
DB 75 ANKHLSTLEDMWNGKDPY 92
```

```
RESULT 13
US-09-499-203-2
; Sequence 2, Application US/09499203
; Patent No. 6570065
; GENERAL INFORMATION:
; APPLICANT: KOSSMANN, Jens
; APPLICANT: WELSH, Thomas
; APPLICANT: QIANZ, Martin
; APPLICANT: KNOTH, Karola
; TITLE OF INVENTION: Nucleic Acid Molecules Encoding Alternansucrase
; FILE REFERENCE: 147-196P
; CURRENT APPLICATION NUMBER: US/09/499,203
; CURRENT FILING DATE: 2000-02-08
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 2057
; TYPE: PRT
; ORGANISM: Leuconostoc mesenteroides
US-09-499-203-2
```

```
Query Match          55.2%; Score 64; DB 4; Length 2057;
Best Local Similarity 66.7%; Pred. No. 0.086;
Matches 12; Conservative 1; Mismatches 5; Indels 0; Gaps 0;
```

```
QY 1 AINHLSTLEAMSDNDPOY 18
DB 665 ANKHLSTLEDMWNGKDPY 682
```

```
RESULT 14
US-09-604-957-7
; Sequence 7, Application US/09604957
; Patent No. 6486314
; GENERAL INFORMATION:
```


GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using SW model

Run on: November 13, 2003, 09:45:40 ; Search time 27.4692 Seconds
(without alignments)
139.565 Million cell updates/sec

Title: US-09-290-049a-11
Perfect score: 116
Sequence: 1 AINHLSTLEAWSNDNDPQYNKD 21

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 666188 seqs, 182559486 residues
Total number of hits satisfying chosen parameters: 666188

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Published Applications AA:*

- 1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
- 2: /cgn2_6/ptodata/2/pubpaa/ECT_NEW_PUB.pep:*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
- 6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep:*
- 7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
- 9: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep:*
- 10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep:*
- 11: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
- 12: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep:*
- 13: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep:*
- 14: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep:*
- 15: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
- 16: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
- 17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
- 18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	116	100.0	545	10 US-09-995-749A-10	Sequence 10, Appl
2	116	100.0	1430	9 US-09-740-274-6	Sequence 6, Appl
3	75	64.7	1475	9 US-09-740-274-2	Sequence 2, Appl
4	68	58.6	522	10 US-09-995-749A-11	Sequence 11, Appl
5	66	55.9	1375	9 US-09-740-274-4	Sequence 4, Appl
6	64	55.2	584	10 US-09-995-749A-12	Sequence 12, Appl
7	53.5	46.1	535	10 US-09-995-749A-13	Sequence 13, Appl
8	53.5	46.1	1781	10 US-09-995-749A-2	Sequence 2, Appl
9	51	44.0	429	9 US-09-815-242-14022	Sequence 14022, A
10	49	42.2	268	15 US-10-156-761-12823	Sequence 12823, A
11	49	42.2	623	12 US-09-933-767-482	Sequence 482, App
12	49	42.2	623	15 US-10-023-282-482	Sequence 482, App
13	49	42.2	704	12 US-09-933-767-422	Sequence 422, App
14	49	42.2	704	15 US-10-023-282-422	Sequence 422, App
15	48	41.4	216	15 US-10-156-761-9158	Sequence 9158, Ap

	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45
	47	46	46	45	44.5	44.5	44	44	43.5	43.5	43.5	43.5	43.5	43.5	43.5	43.5	43.5	43.5	43.5	43.5	43.5	43.5	43.5	43.5	43.5	43.5	43.5	43.5	43.5	43.5
	95	647	703	657	114	1056	212	300	91	192	192	192	192	192	192	192	192	192	192	192	192	192	192	192	192	192	192	192	192	192
	US-09-945-301-14	US-10-116-519-10	US-09-768-877-27	US-09-738-626-6670	US-10-156-761-12167	US-10-098-871-65	US-10-156-761-7969	US-09-794-960-5	US-10-098-871-75	US-09-796-692-1790	US-09-796-692-2054	US-10-040-862-1790	US-10-040-862-2054	US-09-935-868-20	US-10-282-162-20	US-10-282-162-20	US-10-282-162-20	US-10-282-162-20	US-10-282-162-20	US-10-282-162-20	US-10-282-162-20	US-10-282-162-20	US-10-282-162-20	US-10-282-162-20	US-10-282-162-20	US-10-282-162-20	US-10-282-162-20	US-10-282-162-20	US-10-282-162-20	US-10-282-162-20
	Sequence 14, Appl	Sequence 10, Appl	Sequence 27, Appl	Sequence 6670, Ap	Sequence 12167, A	Sequence 65, Appl	Sequence 7969, Ap	Sequence 5, Appl	Sequence 75, Appl	Sequence 1790, Ap	Sequence 2054, Ap	Sequence 1790, Ap	Sequence 2054, Ap	Sequence 20, Appl	Sequence 20, Appl	Sequence 20, Appl	Sequence 20, Appl	Sequence 20, Appl	Sequence 20, Appl	Sequence 20, Appl	Sequence 20, Appl	Sequence 20, Appl	Sequence 20, Appl	Sequence 20, Appl	Sequence 20, Appl	Sequence 20, Appl	Sequence 20, Appl	Sequence 20, Appl	Sequence 20, Appl	Sequence 20, Appl

ALIGNMENTS

RESULT 1
US-09-995-749A-10
Sequence 10, Application US/09995749A
Patent No. US20020155568A1
GENERAL INFORMATION:
APPLICANT: VAN GEL-SCHUTTEN, GERRITDINA HENDRIKA
APPLICANT: DIKHUIZEN, LUBBERT
APPLICANT: RAHAOUI, HAKIM
APPLICANT: LEER, ROBERT-JAN
TITLE OF INVENTION: NOVEL GLUCOSYLTRANSFERASES
FILE REFERENCE: B043388-CIP
CURRENT APPLICATION NUMBER: US/09/995,749A
CURRENT FILING DATE: 2001-11-29
PRIOR APPLICATION NUMBER: 09/604,957
PRIOR FILING DATE: 2000-06-28
PRIOR APPLICATION NUMBER: EPO 00201871.1
PRIOR FILING DATE: 2000-05-25
NUMBER OF SEQ ID NOS: 19
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 10
LENGTH: 545
TYPE: PRT
ORGANISM: Streptococcus mutans
US-09-995-749A-10
Query Match 100.0%; Score 116; DB 10;
Best Local Similarity 100.0%; Pred. No. 1.6e-09;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 AINHLSTLEAWSNDNDPQYNKD 21
DB 75 AINHLSTLEAWSNDNDPQYNKD 95
RESULT 2
US-09-740-274-6
Sequence 6, Application US/09740274
Patent No. US20020031826A1

```
GENERAL INFORMATION:
APPLICANT: Nichols, Scott E.
TITLE OF INVENTION: Glucan-containing Compositions and Paper
FILE REFERENCE: 0357CRD
CURRENT FILING DATE: 2000-12-19
PRIOR APPLICATION NUMBER: 09/210,361
PRIOR FILING DATE: 1998-12-11
PRIOR APPLICATION NUMBER: 09/007,999
PRIOR FILING DATE: 1998-01-16
PRIOR APPLICATION NUMBER: 08/478,704
PRIOR FILING DATE: 1995-06-07
PRIOR APPLICATION NUMBER: 09/009,620
PRIOR FILING DATE: 1998-01-20
PRIOR APPLICATION NUMBER: 08/485,243
PRIOR FILING DATE: 1995-06-07
PRIOR APPLICATION NUMBER: 09/008,172
PRIOR FILING DATE: 1998-01-16
PRIOR APPLICATION NUMBER: 08/482,711
PRIOR FILING DATE: 1995-06-07
NUMBER OF SEQ ID NOS: 6
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 6
LENGTH: 1430
TYPE: PRT
ORGANISM: streptococcus mutans
US-09-740-274-6

Query Match          100.0%; Score 116; DB 9; Length 1430;
Best Local Similarity 100.0%; Pred. No. 4.5e-09;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AINHLSTLEAWSNDNDPOYNKD 21
DB 495 AINHLSTLEAWSNDNDPOYNKD 515

RESULT 3
US-09-740-274-2
Sequence 2, Application US/09740274
Patent No. US20020031826A1
GENERAL INFORMATION:
APPLICANT: Nichols, Scott E.
TITLE OF INVENTION: Glucan-containing Compositions and Paper
FILE REFERENCE: 0357CRD
CURRENT FILING DATE: 2000-12-19
PRIOR APPLICATION NUMBER: 09/210,361
PRIOR FILING DATE: 1998-12-11
PRIOR APPLICATION NUMBER: 09/007,999
PRIOR FILING DATE: 1998-01-16
PRIOR APPLICATION NUMBER: 08/478,704
PRIOR FILING DATE: 1995-06-07
PRIOR APPLICATION NUMBER: 09/009,620
PRIOR FILING DATE: 1998-01-20
PRIOR APPLICATION NUMBER: 08/485,243
PRIOR FILING DATE: 1995-06-07
PRIOR APPLICATION NUMBER: 09/008,172
PRIOR FILING DATE: 1998-01-16
PRIOR APPLICATION NUMBER: 08/482,711
PRIOR FILING DATE: 1995-06-07
NUMBER OF SEQ ID NOS: 6
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 2
LENGTH: 1475
TYPE: PRT
ORGANISM: Streptococcus mutans
US-09-740-274-2

Query Match          64.7%; Score 75; DB 9; Length 1475;
Best Local Similarity 71.4%; Pred. No. 0.0091;
Matches 15; Conservative 1; Mismatches 5; Indels 0; Gaps 0;
```

```
QY 1 AINHLSTLEAWSNDNDPOYNKD 21
DB 481 ANQHLSTLEAWSNDNDPTPLHD 501

RESULT 4
US-09-995-749a-11
Sequence 11, Application US/09995749A
Patent No. US2002015568A1
GENERAL INFORMATION:
APPLICANT: VAN GEL-SCHUTTEN, GERITDINA HENDRIKA
APPLICANT: DIJKHUIZEN, LUBBERT
APPLICANT: RAHMOUI, HAKIM
APPLICANT: LEER, ROBERT-JAN
TITLE OF INVENTION: NOVEL GLUCOSYLTRANSFERASES
FILE REFERENCE: B043388-CIP
CURRENT APPLICATION NUMBER: US/09/995,749A
CURRENT FILING DATE: 2001-11-29
PRIOR APPLICATION NUMBER: 09/604,957
PRIOR FILING DATE: 2000-06-28
PRIOR APPLICATION NUMBER: EPO 00201871.1
PRIOR FILING DATE: 2000-05-25
NUMBER OF SEQ ID NOS: 19
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 11
LENGTH: 522
TYPE: PRT
ORGANISM: Leuconostoc mesenteroides
US-09-995-749a-11

Query Match          58.6%; Score 68; DB 10; Length 522;
Best Local Similarity 66.7%; Pred. No. 0.035;
Matches 14; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 AINHLSTLEAWSNDNDPOYNKD 21
DB 75 ANQHLSTLEAWSNDNDPLVYTD 95

RESULT 5
US-09-740-274-4
Sequence 4, Application US/09740274
Patent No. US20020031826A1
GENERAL INFORMATION:
APPLICANT: Nichols, Scott E.
TITLE OF INVENTION: Glucan-containing Compositions and Paper
FILE REFERENCE: 0357CRD
CURRENT FILING DATE: 2000-12-19
PRIOR APPLICATION NUMBER: 09/210,361
PRIOR FILING DATE: 1998-12-11
PRIOR APPLICATION NUMBER: 09/007,999
PRIOR FILING DATE: 1998-01-16
PRIOR APPLICATION NUMBER: 08/478,704
PRIOR FILING DATE: 1995-06-07
PRIOR APPLICATION NUMBER: 09/009,620
PRIOR FILING DATE: 1998-01-20
PRIOR APPLICATION NUMBER: 08/485,243
PRIOR FILING DATE: 1995-06-07
PRIOR APPLICATION NUMBER: 09/008,172
PRIOR FILING DATE: 1998-01-16
PRIOR APPLICATION NUMBER: 08/482,711
PRIOR FILING DATE: 1995-06-07
NUMBER OF SEQ ID NOS: 6
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 4
LENGTH: 1375
TYPE: PRT
ORGANISM: streptococcus mutans
US-09-740-274-4

Query Match          56.9%; Score 66; DB 9; Length 1375;
Best Local Similarity 66.7%; Pred. No. 0.2;
```

Matches 14; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 1 AINHSLTEAMSDNDPOYKD 21

Db 507 ANHSLTEAMSYNDTPYLDH 527

RESULT 6

US-09-995-749A-12

Sequence 12, Application US/09995749A

Patent No. US20020155568A1

GENERAL INFORMATION:

APPLICANT: VAN GEEL-SCHUTTEN, GERRITDINA HENDRIKA

APPLICANT: DIKHUIZEN, LUBBERT

APPLICANT: RAHAOUI, HAKIM

APPLICANT: LEER, ROBERT-JAN

TITLE OF INVENTION: NOVEL GLUCOSYLTRANSFERASES

FILE REFERENCE: BO43388-CIP

CURRENT APPLICATION NUMBER: US/09/995,749A

PRIOR FILING DATE: 2001-11-29

PRIOR APPLICATION NUMBER: 09/604,957

PRIOR FILING DATE: 2000-06-28

PRIOR APPLICATION NUMBER: EPO 00201871.1

NUMBER OF SEQ ID NOS: 19

SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 12

LENGTH: 584

TYPE: PRT

ORGANISM: Leuconostoc mesenteroides

US-09-995-749A-12

QY 1 AINHSLTEAMSDNDPOY 18

Db 75 ANKHSLTEAMSDNDPOY 92

RESULT 7

US-09-995-749A-13

Sequence 13, Application US/09995749A

Patent No. US20020155568A1

GENERAL INFORMATION:

APPLICANT: VAN GEEL-SCHUTTEN, GERRITDINA HENDRIKA

APPLICANT: DIKHUIZEN, LUBBERT

APPLICANT: RAHAOUI, HAKIM

APPLICANT: LEER, ROBERT-JAN

TITLE OF INVENTION: NOVEL GLUCOSYLTRANSFERASES

FILE REFERENCE: BO43388-CIP

CURRENT APPLICATION NUMBER: US/09/995,749A

PRIOR FILING DATE: 2001-11-29

PRIOR APPLICATION NUMBER: 09/604,957

PRIOR FILING DATE: 2000-06-28

PRIOR APPLICATION NUMBER: EPO 00201871.1

NUMBER OF SEQ ID NOS: 19

SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 13

LENGTH: 535

TYPE: PRT

ORGANISM: Lactobacillus reuteri

US-09-995-749A-13

QY 4 HSLTEAMSDNDPOY-NK 20

Db 77 HINILEDNMHADPEYFNK 94

RESULT 8

US-09-995-749A-2

Sequence 2, Application US/09995749A

Patent No. US20020155568A1

GENERAL INFORMATION:

APPLICANT: VAN GEEL-SCHUTTEN, GERRITDINA HENDRIKA

APPLICANT: DIKHUIZEN, LUBBERT

APPLICANT: RAHAOUI, HAKIM

APPLICANT: LEER, ROBERT-JAN

TITLE OF INVENTION: NOVEL GLUCOSYLTRANSFERASES

FILE REFERENCE: BO43388-CIP

CURRENT APPLICATION NUMBER: US/09/995,749A

PRIOR FILING DATE: 2001-11-29

PRIOR APPLICATION NUMBER: 09/604,957

PRIOR FILING DATE: 2000-06-28

PRIOR APPLICATION NUMBER: EPO 00201871.1

NUMBER OF SEQ ID NOS: 19

SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 2

LENGTH: 1781

TYPE: PRT

ORGANISM: Lactobacillus reuteri

US-09-995-749A-2

QY 4 HSLTEAMSDNDPOY-NK 20

Db 1056 HINILEDNMHADPEYFNK 1073

RESULT 9

US-09-815-242-14022

Sequence 14022, Application US/09815242

Patent No. US20020061569A1

GENERAL INFORMATION:

APPLICANT: Haselbeck, Robert

APPLICANT: Ohlsen, Kari L.

APPLICANT: Zysek, Judith W.

APPLICANT: Wall, Daniel

APPLICANT: Trawick, John D.

APPLICANT: Carr, Grant J.

APPLICANT: Yamamoto, Robert T.

APPLICANT: Xu, H. Howard

TITLE OF INVENTION: Identification of Essential Genes in

FILE REFERENCE: ELITRA.011A

CURRENT APPLICATION NUMBER: US/09/815,242

PRIOR FILING DATE: 2001-03-21

PRIOR APPLICATION NUMBER: 60/191,078

PRIOR FILING DATE: 2000-03-21

PRIOR APPLICATION NUMBER: 60/206,848

PRIOR FILING DATE: 2000-05-23

PRIOR APPLICATION NUMBER: 60/207,727

PRIOR FILING DATE: 2000-05-26

PRIOR APPLICATION NUMBER: 60/242,578

PRIOR FILING DATE: 2000-10-23

PRIOR APPLICATION NUMBER: 60/253,625

PRIOR FILING DATE: 2000-11-27

PRIOR APPLICATION NUMBER: 60/257,931

PRIOR FILING DATE: 2000-12-22

PRIOR APPLICATION NUMBER: 60/269,308

PRIOR FILING DATE: 2001-02-16

NUMBER OF SEQ ID NOS: 14110

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 14022

LENGTH: 429

TYPE: PRT


```

PRIOR APPLICATION NUMBER: 60/070,923
PRIOR FILING DATE: 1997-12-18
PRIOR APPLICATION NUMBER: 60/073,160
PRIOR FILING DATE: 1998-01-30
PRIOR APPLICATION NUMBER: 60/073,159
PRIOR FILING DATE: 1998-01-30
PRIOR APPLICATION NUMBER: 60/073,165
PRIOR FILING DATE: 1998-01-30
PRIOR APPLICATION NUMBER: 60/073,164
PRIOR FILING DATE: 1998-01-30
PRIOR APPLICATION NUMBER: 60/085,925
PRIOR FILING DATE: 1998-05-18
PRIOR APPLICATION NUMBER: 60/085,921
PRIOR FILING DATE: 1998-05-18
PRIOR APPLICATION NUMBER: 60/085,923
PRIOR FILING DATE: 1998-05-18
PRIOR APPLICATION NUMBER: 60/085,922
PRIOR FILING DATE: 1998-05-18
PRIOR APPLICATION NUMBER: 60/092,921
PRIOR FILING DATE: 1998-07-15
PRIOR APPLICATION NUMBER: 60/094,657
PRIOR FILING DATE: 1998-07-30
NUMBER OF SEQ ID NOS: 1245
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 482
LENGTH: 623
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SITE
LOCATION: (111)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (575)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-933-767-482

Query March 42.2% Score 49; DB 12; Length 623;
Best Local Similarity 52.9%; Pred. No. 35;
Matches 9; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 5 LSLFMSDNDPOYNKD 21
Db 107 IMVLXGSEFDPQYNKD 123

RESULT 12
US-10-023-282-482
Sequence 482, Application US/10023282
Publication No. US20030092893A1
GENERAL INFORMATION:
APPLICANT: Young et al.
TITLE OF INVENTION: 207 Human Secreted Proteins
FILE REFERENCE: P2007P1
CURRENT APPLICATION NUMBER: US/10/023,282
EARLIER FILING DATE: 2001-12-20
EARLIER APPLICATION NUMBER: 09/205,258
EARLIER FILING DATE: 1998-12-04
EARLIER APPLICATION NUMBER: PCT/US98/11422
EARLIER FILING DATE: 1998-06-04
EARLIER APPLICATION NUMBER: 60/048,885
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/049,375
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,881
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,880
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,896
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/049,020
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,876
```

```

EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,895
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,884
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,894
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,971
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,964
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,882
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,899
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,893
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,900
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,901
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,892
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,915
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/049,019
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,970
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,972
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,916
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/049,373
EARLIER FILING DATE: 1997-06-05
EARLIER APPLICATION NUMBER: 60/048,875
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/049,374
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,917
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,949
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,974
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,883
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,897
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,898
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,962
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,963
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,877
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,878
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/070,923
EARLIER FILING DATE: 1997-12-18
EARLIER APPLICATION NUMBER: 60/092,921
EARLIER FILING DATE: 1998-07-15
EARLIER APPLICATION NUMBER: 60/094,657
EARLIER FILING DATE: 1998-07-30
NUMBER OF SEQ ID NOS: 1227
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 482
LENGTH: 623
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SITE
```

```
LOCATION: (111)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
FEATURE:
NAME/KEY: SITE
LOCATION: (575)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-10-023-282-482

Query Match      42.2%; Score 49; DB 15; Length 623;
Best Local Similarity 52.9%; Pred. No. 35;
Matches 9; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

Cy      5 ISLEAWSNDPOXNKD 21
Db      107 IMVLXGASEPDPQNKD 123

RESULT 13
US-09-933-767-422
Sequence 422, Application US/09933767
Publication No. US20030181692A1
GENERAL INFORMATION:
APPLICANT: Ni et al.
TITLE OF INVENTION: 207 Human Secreted Proteins
FILE REFERENCE: P2007P2
CURRENT APPLICATION NUMBER: US/09/933,767
PRIOR FILING DATE: 2001-02-21
PRIOR APPLICATION NUMBER: 60/184,836
PRIOR FILING DATE: 2000-02-24
PRIOR APPLICATION NUMBER: 60/193,170
PRIOR FILING DATE: 2000-03-29
PRIOR APPLICATION NUMBER: 09/205,258
PRIOR FILING DATE: 1998-12-04
PRIOR APPLICATION NUMBER: PCT/US98/11422
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/048,885
PRIOR FILING DATE: 1997-06-06
PRIOR APPLICATION NUMBER: 60/049,375
PRIOR FILING DATE: 1997-06-06
PRIOR APPLICATION NUMBER: 60/048,881
PRIOR FILING DATE: 1997-06-06
PRIOR APPLICATION NUMBER: 60/048,880
PRIOR FILING DATE: 1997-06-06
PRIOR APPLICATION NUMBER: 60/048,896
PRIOR FILING DATE: 1997-06-06
PRIOR APPLICATION NUMBER: 60/049,020
PRIOR FILING DATE: 1997-06-06
PRIOR APPLICATION NUMBER: 60/048,876
PRIOR FILING DATE: 1997-06-06
PRIOR APPLICATION NUMBER: 60/048,895
PRIOR FILING DATE: 1997-06-06
PRIOR APPLICATION NUMBER: 60/048,884
PRIOR FILING DATE: 1997-06-06
PRIOR APPLICATION NUMBER: 60/048,894
PRIOR FILING DATE: 1997-06-06
PRIOR APPLICATION NUMBER: 60/048,971
PRIOR FILING DATE: 1997-06-06
PRIOR APPLICATION NUMBER: 60/048,964
PRIOR FILING DATE: 1997-06-06
PRIOR APPLICATION NUMBER: 60/048,882
PRIOR FILING DATE: 1997-06-06
PRIOR APPLICATION NUMBER: 60/048,899
PRIOR FILING DATE: 1997-06-06
PRIOR APPLICATION NUMBER: 60/048,893
PRIOR FILING DATE: 1997-06-06
PRIOR APPLICATION NUMBER: 60/048,900
PRIOR FILING DATE: 1997-06-06
PRIOR APPLICATION NUMBER: 60/048,901
PRIOR FILING DATE: 1997-06-06
PRIOR APPLICATION NUMBER: 60/048,892
PRIOR FILING DATE: 1997-06-06
```

```
PRIOR APPLICATION NUMBER: 60/048,915
PRIOR FILING DATE: 1997-06-06
PRIOR APPLICATION NUMBER: 60/049,019
PRIOR FILING DATE: 1997-06-06
PRIOR APPLICATION NUMBER: 60/048,970
PRIOR FILING DATE: 1997-06-06
PRIOR APPLICATION NUMBER: 60/048,972
PRIOR FILING DATE: 1997-06-06
PRIOR APPLICATION NUMBER: 60/048,916
PRIOR FILING DATE: 1997-06-06
PRIOR APPLICATION NUMBER: 60/049,373
PRIOR FILING DATE: 1997-06-06
PRIOR APPLICATION NUMBER: 60/048,875
PRIOR FILING DATE: 1997-06-06
PRIOR APPLICATION NUMBER: 60/049,374
PRIOR FILING DATE: 1997-06-06
PRIOR APPLICATION NUMBER: 60/048,917
PRIOR FILING DATE: 1997-06-06
PRIOR APPLICATION NUMBER: 60/048,949
PRIOR FILING DATE: 1997-06-06
PRIOR APPLICATION NUMBER: 60/048,974
PRIOR FILING DATE: 1997-06-06
PRIOR APPLICATION NUMBER: 60/048,883
PRIOR FILING DATE: 1997-06-06
PRIOR APPLICATION NUMBER: 60/048,897
PRIOR FILING DATE: 1997-06-06
PRIOR APPLICATION NUMBER: 60/048,898
PRIOR FILING DATE: 1997-06-06
PRIOR APPLICATION NUMBER: 60/048,962
PRIOR FILING DATE: 1997-06-06
PRIOR APPLICATION NUMBER: 60/048,963
PRIOR FILING DATE: 1997-06-06
PRIOR APPLICATION NUMBER: 60/048,877
PRIOR FILING DATE: 1997-06-06
PRIOR APPLICATION NUMBER: 60/048,878
PRIOR FILING DATE: 1997-06-06
PRIOR APPLICATION NUMBER: 60/068,054
PRIOR FILING DATE: 1997-12-18
PRIOR APPLICATION NUMBER: 60/068,064
PRIOR FILING DATE: 1997-12-18
PRIOR APPLICATION NUMBER: 60/068,053
PRIOR FILING DATE: 1997-12-18
PRIOR APPLICATION NUMBER: 60/070,923
PRIOR FILING DATE: 1997-12-18
PRIOR APPLICATION NUMBER: 60/073,160
PRIOR FILING DATE: 1998-01-30
PRIOR APPLICATION NUMBER: 60/073,159
PRIOR FILING DATE: 1998-01-30
PRIOR APPLICATION NUMBER: 60/073,165
PRIOR FILING DATE: 1998-01-30
PRIOR APPLICATION NUMBER: 60/073,164
PRIOR FILING DATE: 1998-01-30
PRIOR APPLICATION NUMBER: 60/085,925
PRIOR FILING DATE: 1998-05-18
PRIOR APPLICATION NUMBER: 60/085,921
PRIOR FILING DATE: 1998-05-18
PRIOR APPLICATION NUMBER: 60/085,923
PRIOR FILING DATE: 1998-05-18
PRIOR APPLICATION NUMBER: 60/085,922
PRIOR FILING DATE: 1998-05-18
PRIOR APPLICATION NUMBER: 60/092,921
PRIOR FILING DATE: 1998-07-15
PRIOR APPLICATION NUMBER: 60/094,657
PRIOR FILING DATE: 1998-07-30
NUMBER OF SEQ ID NOS: 1245
SOFTWARE: Patentn Ver. 2.0
SEQ ID NO 422
LENGTH: 704
TYPE: PRT
ORGANISM: Homo sapiens
US-09-933-767-422

Query Match      42.2%; Score 49; DB 12; Length 704;
```

Best Local Similarity 52.9%; Pred. No. 40;
Matches 9; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 5 LSTLEAWSNDPOYNKD 21
Db 191 IMVLGASFPDPYTKD 207

RESULT 14
US-10-023-282-422
; Sequence 422, Application US/10023282
; Publication No. US20030092893A1
; GENERAL INFORMATION:
; APPLICANT: Young et al.
; TITLE OF INVENTION: 207 Human Secreted Proteins
; FILE REFERENCE: P2007P1
; CURRENT APPLICATION NUMBER: US/10/023,282
; CURRENT FILING DATE: 2001-12-20
; EARLIER APPLICATION NUMBER: 09/205,258
; EARLIER FILING DATE: 1998-12-04
; EARLIER APPLICATION NUMBER: PCT/US98/11422
; EARLIER FILING DATE: 1998-06-04
; EARLIER APPLICATION NUMBER: 60/048,885
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,375
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,881
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,880
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,896
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,020
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,876
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,895
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,884
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,894
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,971
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,964
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,882
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,899
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,893
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,900
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,901
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,892
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,915
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,019
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,970
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,972
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,916
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,373
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,875
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,374

; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,917
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,949
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,974
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,883
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,897
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,898
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,962
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,963
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,877
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,878
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/070,923
; EARLIER FILING DATE: 1997-12-18
; EARLIER APPLICATION NUMBER: 60/092,921
; EARLIER FILING DATE: 1998-07-15
; EARLIER APPLICATION NUMBER: 60/094,657
; EARLIER FILING DATE: 1998-07-30
; NUMBER OF SEQ ID NOS: 1227
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 422
; LENGTH: 704
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-023-282-422

Query Match 42.2%; Score 49; DB 15; Length 704;
Best Local Similarity 52.9%; Pred. No. 40;
Matches 9; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 5 LSTLEAWSNDPOYNKD 21
Db 191 IMVLGASFPDPYTKD 207

RESULT 15
US-10-156-761-9158
; Sequence 9158, Application US/10156761
; Publication No. US20030115018A1
; GENERAL INFORMATION:
; APPLICANT: OMDRA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 9158
; LENGTH: 216
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-9158

Query Match 41.4%; Score 48; DB 15; Length 216;
Best Local Similarity 64.3%; Pred. No. 16;

Matches 9; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 4 HLSTLEAWSNDPQ 17
|||::|||

Db 184 HLSMTEAIRDRDPQ 197

Search completed: November 13, 2003, 10:29:02
Job time : 28.4692 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 13, 2003, 09:31:40 ; Search time 38.5166 Seconds
(without alignments)
86.541 Million cell updates/sec

Title: US-09-290-049A-12
Perfect score: 119
Sequence: 1 ANNHVSIVBAMSDNTPYLD 21

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues
Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

1: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
6: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
7: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
8: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
9: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
10: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
11: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
12: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
13: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
14: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
15: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
16: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
17: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
18: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
19: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
20: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
21: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
22: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*
23: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.*
24: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	119	100.0	1592	AA32925	Glucosyltransferase
2	108	90.8	1017	AAU79285	Streptococcus muta
3	108	90.8	1475	AAU98027	S. mutans glucosyl
4	108	90.8	1475	AAU98030	S. mutans glucosyl
5	108	90.8	1475	AAU98031	S. mutans glucosyl
6	108	90.8	1475	AAU98032	S. mutans glucosyl
7	108	90.8	1475	AAU98033	S. mutans glucosyl
8	108	90.8	1475	AAU98034	S. mutans glucosyl
9	108	90.8	1475	AAU98035	S. mutans glucosyl

10	108	90.8	1475	AAU98035	S. mutans glucosyl
11	108	90.8	1475	AAU98037	S. mutans glucosyl
12	108	90.8	1475	AAU98038	S. mutans GTPB mut
13	108	90.8	1475	AAU98039	S. mutans glucosyl
14	108	90.8	1475	AAU98040	S. mutans glucosyl
15	108	90.8	1475	AAU98041	Streptococcus muta
16	99	83.2	1375	AAU98028	S. mutans glucosyl
17	99	83.2	1375	AAU98029	Streptococcus muta
18	99	83.2	1375	AAU98028	S. mutans glucosyl
19	74	62.2	1430	AAU98029	S. mutans glucosyl
20	74	62.2	1430	AAU98042	S. mutans glucosyl
21	74	62.2	1430	AAU98043	S. mutans glucosyl
22	74	62.2	1430	AAU98044	S. mutans glucosyl
23	74	62.2	1430	AAU98045	S. mutans glucosyl
24	72	60.5	2835	ABR98574	Dextran-saccharase
25	66	55.5	12	ABR98642	Dextran-saccharase
26	65	54.6	1527	AAU80055	Leuconostoc mesent
27	60	50.4	12	ABR98641	Dextran-saccharase
28	57	47.9	1577	AAU91047	Alpha-D-glucosyltr
29	55	46.2	2057	ABR10667	L. mesenteroides a
30	53	44.5	12	ABR98643	Dextran-saccharase
31	53	44.5	195	AAU09941	Arabidopsis thalia
32	52	43.7	855	ABR98573	Dextran-saccharase
33	51	42.9	1195	AAU75420	T. thermophilus nit
34	50	42.0	1252	ABR59363	Drosophila melanog
35	48.5	40.8	401	AAU43544	Human polypeptide
36	48	40.3	1781	AAU74519	Lactobacillus reut
37	47	39.5	12	ABR98578	Dextran-saccharase
38	47	39.5	12	ABR98644	Dextran-saccharase
39	47	39.5	12	ABR98646	Dextran-saccharase
40	47	39.5	12	ABR98648	Dextran-saccharase
41	47	39.5	749	AAU87244	LDL receptor bindi
42	46.5	39.1	539	AAU87244	Novel central nerv
43	46.5	39.1	1604	AAU82715	Amino acid sequenc
44	46	38.7	486	AAU36552	Lawsonia intracell
45	46	38.7	880	ABR61809	Drosophila melanog

ALIGNMENTS

RESULT 1	AA32925
ID	AA32925 standard; Protein; 1592 AA.
XX	AA32925;
AC	28-JUN-1993 (first entry)
XX	
DT	Glucosyltransferase I.
XX	
DE	GT-1; Streptococcus; dental; carries.
XX	
KW	Streptococcus sobrinus.
XX	
OS	JP05023188-A.
XX	
FN	02-FEB-1993.
XX	
PD	25-JUL-1991; 91UP-0186592.
XX	
PF	25-JUL-1991; 91UP-0186552.
XX	
PR	(FUKU//) FUKU I.
XX	(KATO//) KATO K.
PA	WPI: 1993-079449/10.
XX	N-PSDB; AAQ37760.
DR	DNA sequence glucosyl:transferase-I - comprises Streptococcus
XX	sobrinus DNA sequence with at least one nucleotide added or
PT	deleted
XX	

PS Claim 13; Page 15; 29pp; Japanese.
 XX
 CC The DNA sequence from *Streptococcus sobrinus* strain 6715 encodes
 CC glucosyltransferase-I (and mutants). The DNA was obtd. by treating
 CC *S. sobrinus* 6715 with mutanolysin, extracting the chromosomal DNA,
 CC partially digesting with *Sau3AI* and fractionating on agarose gel.
 CC The 3-5 kbp fragment was ligated into pUC18 and *E. coli* JM109
 CC transformed with it. A *Gf*-1 expressing clone was isolated and
 CC sequenced. The clone may be used in the development of a drug for
 CC dental caries.
 CC
 XX Sequence 1592 AA;
 SQ
 Query Match 100.0%; Score 119; DB 14; Length 1592;
 Best Local Similarity 100.0%; Pred. No. 6.4e-09;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 ANNHVSIVEAMSDNDTPYLHD 21
 DB 477 ANNHVSIVEAMSDNDTPYLHD 497
 RESULT 2
 AAU79285
 ID AAU79285 standard; Protein; 1017 AA.
 XX
 AC AAU79285;
 XX
 DT 13-AUG-2002 (first entry)
 XX
 DE Streptococcus mutans monoclonal antibody-related protein #2.
 XX
 KW Antibody; dental caries; water insoluble glucan synthetase;
 KW anti-carries; glucosyl transferase-B; immunotherapy.
 XX
 OS Streptococcus mutans.
 XX
 PN JP2002114709-A.
 XX
 PD 16-APR-2002.
 XX
 PF 04-OCT-2000; 2000JP-0304893.
 XX
 PR 04-OCT-2000; 2000JP-0304893.
 XX
 PA (UNIV-) UNIV NIPPON.
 XX
 DR WPI; 2002-448101/48.
 XX
 PT Anti-carries agent composed of a monoclonal antibody against an
 PT inhibitory enzyme against water insoluble glucan synthetase of glucosyl
 PT transferase-B (*Gf*-B) of *Streptococcus mutans* -
 XX
 PS Claim 4; Page 17-19; 28pp; Japanese.
 XX
 CC The invention relates to a monoclonal antibody against dental caries and
 CC an anti-carries agent composed of a monoclonal antibody produced by
 CC *Streptococcus mutans*, particularly mouse-hybridoma MHP126 (FERM P-17566)
 CC or mouse-hybridoma MHP136 (FERM P-17567), against an enzyme having
 CC inhibitive activity against water insoluble glucan synthetase of glucosyl
 CC transferase-B. The monoclonal antibody specifically inhibits water
 CC insoluble glucan synthetase of *Streptococcus mutans* produced glucosyl
 CC transferase-B and is used in the immunotherapy of dental caries. This
 CC sequence represents a *Streptococcus mutans* monoclonal antibody-related
 CC protein.
 CC
 XX Sequence 1017 AA;
 SQ
 Query Match 90.8%; Score 108; DB 23; Length 1017;
 Best Local Similarity 85.7%; Pred. No. 1.7e-07;
 Matches 18; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
 QY 1 ANNHVSIVEAMSDNDTPYLHD 21

DB 447 ANNHVSIVEAMSDNDTPYLHD 467
 RESULT 3
 AAU98027
 ID AAU98027 standard; Protein; 1475 AA.
 XX
 AC AAU98027;
 XX
 DT 27-AUG-2002 (first entry)
 XX
 DE S. mutans glucosyltransferase *GTFB*.
 XX
 KW glucosyltransferase; *GTFB*; transgenic plant; paper sizing;
 KW coating composition; glucan; starch; latex; thermoplastic molecule;
 KW amyloplast; vacuole; paper manufacture.
 XX
 OS Streptococcus mutans.
 XX
 PN US2002031826-A1.
 XX
 PD 14-MAR-2002.
 XX
 PF 19-DEC-2000; 2000US-0740274.
 XX
 PR 11-DEC-1998; 98US-0210361.
 PR 07-JUN-1995; 95US-0478704.
 PR 07-JUN-1995; 95US-0482711.
 PR 07-JUN-1995; 95US-0485243.
 PR 16-JAN-1998; 98US-0007999.
 PR 16-JAN-1998; 98US-0008172.
 PR 20-JAN-1998; 98US-0009620.
 XX
 PA (NICH/) NICHOLS S E.
 XX
 PI Nichols SE;
 XX
 DR WPI; 2002-414332/44.
 DR N-PSDB; ABK52938.
 XX
 PT Glucosyltransferase B or D protein useful for producing a glucan useful
 PT as substitutes for and additions to modified starch and latexes in
 PT paper manufacture, comprises mutations in specific positions -
 XX
 PS Disclosure; Page 21-25; 44pp; English.
 XX
 CC The invention an isolated protein comprising a glucosyltransferase
 CC (*GTF*) B polypeptide having changes at position from 1448V, D457N,
 CC D567T, K1014T, D457N/D567T, D457N/D571K, D567T/D571K,
 CC D567T/D571K/K1014T, I448V/D457N/D567T/D571K/K790/K1014T,
 CC Y169A/Y170A/Y171A, and K779Q or a *GTF* D polypeptide having
 CC changes at positions from T589D, T589E, N471D, N471D/T589D, and
 CC N471D/T589E. Also included are a glucan produced by the *GTF* mutant,
 CC an isolated polynucleotide which encodes P1 or P2, or its complementary
 CC polynucleotide, a ribonucleic acid sequence encoding the *GTF* mutant,
 CC an expression cassette comprising the polynucleotide operably linked to a
 CC promoter, a vector comprising the expression cassette, host cell
 CC introduced with the vector, a transgenic plant comprising the
 CC vector, a seed or tuber from the transgenic plant, a paper sizing and/or
 CC coating composition comprising a glucan produced in a plant transformed
 CC with a gene encoding the mutant *GTF*, wild type or, starch, a latex,
 CC thermoplastic molecule or their combinations or glucan and starch where
 CC the glucan is produced in the amyloplast and/or vacuole or a maize line
 CC deficient in starch biosynthesis, transformed with a gene encoding a
 CC glucosyltransferase B or D enzyme, wild-type or mutant and a paper
 CC comprising the glucan (paper sizing/coating agent). The vector is useful
 CC for producing a glucan in a plant. The method comprises transforming a
 CC plant cell with the vector, growing the plant cell under plant growing
 CC conditions to produce a regenerated plant and inducing expression of the
 CC polynucleotide for a time sufficient to produce the glucan in the
 CC regenerated plant, where the vector contains a transit sequence from
 CC ribulose biphosphate carboxylase small subunit, waxy, brittle-1 and

CC chlorophyll AB binding protein to produce a transgenic plant, and glucan
 CC is produced in the amyloplast of potato or the vacuole of sugar beet.
 CC Glucans are useful as substitutes for and additions to modified starch
 CC and latexes in paper manufacture. Unlike prior art techniques, which
 CC require input materials that produce chemical effluents, paper
 CC manufacture utilizing the glucan produced by GTF, which utilizes
 CC biologically produced input materials, is more cost-effective and
 CC environmentally friendly. Moreover, glucans also exhibit thermoplastic
 CC properties and impart gloss to the paper during coating step.
 CC The present sequence represents GTFB.
 XX
 SQ Sequence 1475 AA;
 Query Match 90.8%; Score 108; DB 23; Length 1475;
 Best Local Similarity 85.7%; Pred. NO. 2.7e-07;
 Matches 18; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
 QY 1 ANNHVSIVEAMSDNDPYLHD 21
 DB 481 ANDHLSILEAMSDNDPYLHD 501
 RESULT 4
 AAU98030 standard; Protein: 1475 AA.
 AAU98030;
 27-AUG-2002 (first entry)
 S. mutans glucosyltransferase GTFB mutant I448V.
 XX Glucosyltransferase; GTFB; transgenic plant; paper sizing;
 KW coating composition; glucan; starch; latex; thermoplastic molecule;
 KW amyloplast; vacuole; paper manufacture; mutant; mutcin.
 XX Streptococcus mutans.
 OS Synthetic.
 XX Key Location/Qualifiers
 FT Misc-difference 448 /note= "Wild-type Ile substituted by Val"
 FT
 FT US2002031826-A1.
 PN 14-MAR-2002.
 PD 19-DEC-2000; 2000US-0740274.
 PF 11-DEC-1998; 98US-0210361.
 XX 07-JUN-1995; 95US-0478704.
 PR 07-JUN-1995; 95US-0482711.
 PR 07-JUN-1995; 95US-0485243.
 PR 16-JAN-1998; 98US-0007999.
 PR 16-JAN-1998; 98US-0008172.
 PR 20-JAN-1998; 98US-0009620.
 XX
 XX (NICH) NICHOLS S E.
 PA
 XX Nichols SE;
 PI
 DR WPI, 2002-414332/44.
 XX
 PT Glucosyltransferase B or D protein useful for producing a glucan useful
 PT as substitutes for and additions to modified starch and latexes in
 PT paper manufacture, comprises mutations in specific positions -
 XX
 PS Claim 36; Page -; 44pp; English.
 XX
 CC The invention an isolated protein comprising a glucosyltransferase
 CC (GTF) B polypeptide having changes at position from I448V, D457N,
 CC D567T, K1014T, D457N/D567T, D457N/D571K, D567T/D571K,
 CC D567T/D571K/K1014T, I448V/D457N/D567T/D571K/K779Q/K1014T,

CC Y169A/Y170A/Y171A, and K779Q or a GTF D polypeptide having
 CC changes at positions from T589D, T589E, N471D, N471D/T589D, and
 CC N471D/T589E. Also included are a glucan produced by the GTF mutant,
 CC an isolated polynucleotide which encodes P1 or P2, or its complementary
 CC polynucleotide, a ribonucleic acid sequence encoding the GTF mutant,
 CC an expression cassette comprising the polynucleotide operably linked to a
 CC promoter, a vector comprising the expression cassette, host cell
 CC introduced with the vector, a transgenic plant comprising the
 CC vector, a seed or tuber from the transgenic plant, a paper sizing and/or
 CC coating composition comprising a glucan produced in a plant transformed
 CC with a gene encoding the mutant GTF, wild type or, starch, a latex,
 CC thermoplastic molecule or their combinations or glucan and starch where
 CC the glucan is produced in the amyloplast and/or vacuole or a maize line
 CC deficient in starch biosynthesis, transformed with a gene encoding a
 CC glucosyltransferase B or D enzyme, wild-type or mutant and a paper
 CC comprising the glucan (paper sizing/coating agent). The vector is useful
 CC for producing a glucan in a plant. The method comprises transforming a
 CC plant cell with the vector, growing the plant cell under plant growing
 CC conditions to produce a regenerated plant and inducing expression of the
 CC polynucleotide for a time sufficient to produce the glucan in the
 CC regenerated plant, where the vector contains a transit sequence from
 CC ribulose biphosphate carboxylase small subunit, waxy, brittle-1 and
 CC chlorophyll AB binding protein to produce a transgenic plant, and glucan
 CC is produced in the amyloplast of potato or the vacuole of sugar beet.
 CC Glucans are useful as substitutes for and additions to modified starch
 CC and latexes in paper manufacture. Unlike prior art techniques, which
 CC require input materials that produce chemical effluents, paper
 CC manufacture utilizing the glucan produced by GTF, which utilizes
 CC biologically produced input materials, is more cost-effective and
 CC environmentally friendly. Moreover, glucans also exhibit thermoplastic
 CC properties and impart gloss to the paper during coating step.
 CC The present sequence represents a GTFB mutant of the invention.
 CC Note: The present sequence is not shown in the specification but
 CC was created by the indexer using the GTFB sequence appearing as AAU98027
 CC and the information in claim 36.
 XX
 SQ Sequence 1475 AA;
 Query Match 90.8%; Score 108; DB 23; Length 1475;
 Best Local Similarity 85.7%; Pred. NO. 2.7e-07;
 Matches 18; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
 QY 1 ANNHVSIVEAMSDNDPYLHD 21
 DB 481 ANDHLSILEAMSDNDPYLHD 501

RESULT 5
 AAU98031 standard; Protein: 1475 AA.
 AAU98031;
 27-AUG-2002 (first entry)
 S. mutans glucosyltransferase GTFB mutant D457N.
 XX Glucosyltransferase; GTFB; transgenic plant; paper sizing;
 KW coating composition; glucan; starch; latex; thermoplastic molecule;
 KW amyloplast; vacuole; paper manufacture; mutant; mutcin.
 XX Streptococcus mutans.
 OS Synthetic.
 XX Key Location/Qualifiers
 FT Misc-difference 457 /note= "Wild-type Asp substituted by Asn"
 FT
 FT US2002031826-A1.
 PN 14-MAR-2002.
 PD 19-DEC-2000; 2000US-0740274.
 PF

XX 11-DEC-1998; 98US-0210361.
 PR 07-JUN-1995; 95US-0478704.
 PR 07-JUN-1995; 95US-0482711.
 PR 07-JUN-1995; 95US-0485243.
 PR 16-JAN-1998; 98US-0007999.
 PR 16-JAN-1998; 98US-0008172.
 PR 20-JAN-1998; 98US-0009620.
 XX
 PA (NICH/) NICHOLS S E.
 XX
 PI Nichols SE;
 XX
 DR WPI; 2002-414332/44.
 XX
 PT Glucosyltransferase B or D protein useful for producing a glucan useful
 PT as substitutes for and additions to modified starch and latexes in
 PT paper manufacture, comprises mutations in specific positions -
 XX
 PS Claim 36; Page -; 44pp; English.

CC The invention an isolated protein comprising a glucosyltransferase
 CC (GTF) B polypeptide having changes at position from 1448V, D457N,
 CC D567T, K1014T, D457N/D567T, D457N/D571K, D567T/D571K,
 CC D567T/D571K/K1014T, 1448V/D457N/D567T/D571K/K779Q/K1014T,
 CC Y169A/Y170A/Y171A, and K779Q or a GTF D polypeptide having
 CC changes at positions from T589D, T589E, N471D, N471D/T589D, and
 CC N471D/T589E. Also included are a glucan produced by the GTF mutant,
 CC an isolated polynucleotide which encodes P1 or P2, or its complementary
 CC polynucleotide, a ribonucleic acid sequence encoding the GTF mutant,
 CC an expression cassette comprising the polynucleotide operably linked to a
 CC promoter, a vector comprising the expression cassette, host cell
 CC introduced with the vector, a transgenic plant comprising the
 CC vector, a seed or tuber from the transgenic plant, a paper sizing and/or
 CC coating composition comprising a glucan produced in a plant transformed
 CC with a gene encoding the mutant GTF, wild type or, starch, a latex,
 CC thermoplastic molecule or their combinations or glucan and starch where
 CC the glucan is produced in the amyloplast and/or vacuole or a maize line
 CC deficient in starch biosynthesis, transformed with a gene encoding a
 CC glucosyltransferase B or D enzyme, wild-type or mutant and a paper
 CC comprising the glucan (paper sizing/coating agent). The vector is useful
 CC for producing the glucan in a plant. The method comprises transforming a
 CC plant cell with the vector, growing the plant cell under plant growing
 CC conditions to produce a regenerated plant and inducing expression of the
 CC polynucleotide for a time sufficient to produce the glucan in the
 CC regenerated plant, where the vector contains a transit sequence from
 CC ribulose biphosphate carboxylase small subunit, waxy, brittle-1 and
 CC chlorophyll AB binding protein to produce a transgenic plant, and glucan
 CC is produced in the amyloplast of potato or the vacuole of sugar beet.
 CC Glucans are useful as substitutes for and additions to modified starch
 CC and latexes in paper manufacture. Unlike prior art techniques, which
 CC require input materials that produce chemical effluents, paper
 CC manufacture utilizing the glucan produced by GTF, which utilizes
 CC biologically produced input materials, is more cost-effective and
 CC environmentally friendly. Moreover, glucans also exhibit thermoplastic
 CC properties and impart gloss to the paper during coating step.
 CC The present sequence represents a GTFB mutant of the invention.
 CC Note: The present sequence is not shown in the specification but
 CC was created by the indexer using the GTFB sequence appearing as AAU98027
 CC and the information in claim 36.
 XX
 XX Sequence 1475 AA;
 SQ

Query Match 90.8%; Score 108; DB 23; Length 1475;
 Best Local Similarity 85.7%; Pred. No. 2.7e-07;
 Matches 18; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 ANNHVSIWEASNDPTPYLHD 21
 ||:|||||
 Db 481 ANDHUSILFEASNDPTPYLHD 501

RESULT 6

AAU98032
 ID AAU98032 standard; Protein, 1475 AA.
 XX
 AC AAU98032;
 XX
 DT 27-AUG-2002 (first entry)
 XX
 DE S. mutans glucosyltransferase GTFB mutant D567T.
 XX
 KM Glucosyltransferase; GTFB; transgenic plant; paper sizing;
 KM coating composition; glucan; starch; latex; thermoplastic molecule;
 KM amyloplast; vacuole; paper manufacture; mutant; mutlein.
 XX
 OS Streptococcus mutans.
 OS Synthetic.
 OS
 FT Key Location/Qualifiers
 FT Misc-difference 567 /note= "Wild-type Asp substituted by Thr"
 FT
 XX
 PN US2002031826-A1.
 XX
 PD 14-MAR-2002.
 XX
 PF 19-DEC-2000; 2000US-0740274.
 XX
 XX 11-DEC-1998; 98US-0210361.
 PR 07-JUN-1995; 95US-0478704.
 PR 07-JUN-1995; 95US-0482711.
 PR 07-JUN-1995; 95US-0485243.
 PR 16-JAN-1998; 98US-0007999.
 PR 16-JAN-1998; 98US-0008172.
 PR 20-JAN-1998; 98US-0009620.
 XX
 PA (NICH/) NICHOLS S E.
 XX
 PI Nichols SE;
 XX
 DR WPI; 2002-414332/44.
 XX
 PT Glucosyltransferase B or D protein useful for producing a glucan useful
 PT as substitutes for and additions to modified starch and latexes in
 PT paper manufacture, comprises mutations in specific positions -
 XX
 PS Claim 36; Page -; 44pp; English.

CC The invention an isolated protein comprising a glucosyltransferase
 CC (GTF) B polypeptide having changes at position from 1448V, D457N,
 CC D567T, K1014T, D457N/D567T, D457N/D571K, D567T/D571K,
 CC D567T/D571K/K1014T, 1448V/D457N/D567T/D571K/K779Q/K1014T,
 CC Y169A/Y170A/Y171A, and K779Q or a GTF D polypeptide having
 CC changes at positions from T589D, T589E, N471D, N471D/T589D, and
 CC N471D/T589E. Also included are a glucan produced by the GTF mutant,
 CC an isolated polynucleotide which encodes P1 or P2, or its complementary
 CC polynucleotide, a ribonucleic acid sequence encoding the GTF mutant,
 CC an expression cassette comprising the polynucleotide operably linked to a
 CC promoter, a vector comprising the expression cassette, host cell
 CC introduced with the vector, a transgenic plant comprising the
 CC vector, a seed or tuber from the transgenic plant, a paper sizing and/or
 CC coating composition comprising a glucan produced in a plant transformed
 CC with a gene encoding the mutant GTF, wild type or, starch, a latex,
 CC thermoplastic molecule or their combinations or glucan and starch where
 CC the glucan is produced in the amyloplast and/or vacuole or a maize line
 CC deficient in starch biosynthesis, transformed with a gene encoding a
 CC glucosyltransferase B or D enzyme, wild-type or mutant and a paper
 CC comprising the glucan (paper sizing/coating agent). The vector is useful
 CC for producing the glucan in a plant. The method comprises transforming a
 CC plant cell with the vector, growing the plant cell under plant growing
 CC conditions to produce a regenerated plant and inducing expression of the
 CC polynucleotide for a time sufficient to produce the glucan in the
 CC regenerated plant, where the vector contains a transit sequence from
 CC ribulose biphosphate carboxylase small subunit, waxy, brittle-1 and
 CC chlorophyll AB binding protein to produce a transgenic plant, and glucan

CC	is produced in the amyloplast of potato or the vacuole of sugar beet.
CC	Glucans are useful as substitutes for and additions to modified starch
CC	and latexes in paper manufacture. Unlike prior art techniques, which
CC	require input materials that produce chemical effluents, paper
CC	manufacture utilizing the glucan produced by GTF, which utilizes
CC	biologically produced input materials, is more cost-effective and
CC	environmentally friendly. Moreover, glucans also exhibit thermoplastic
CC	properties and impart gloss to the paper during coating step.
CC	The present sequence represents a GTF mutant of the invention.
CC	Note: The present sequence is not shown in the specification but
CC	was created by the indexer using the GTFB sequence appearing as AAU98027
CC	and the information in claim 36.
XX	
SQ	Sequence 1475 AA;
Query Match	90.8%; Score 108; DB 23; Length 1475;
Best Local Similarity	85.7%; Pred. No. 2.7e-07;
Matches 18; Conservative	3; Mismatches 0; Indels 0; Gaps 0;
QY	1 ANNHVSIVGANSNDTPYLHD 21
	: : : : : : : : :
Db	481 ANDHLSIIEAWSNDTPYLHD 501
RESULT 7	
ID	AAU98033
XX	AAU98033 standard; Protein; 1475 AA.
AC	AAU98033;
XX	
DT	27-AUG-2002 (first entry)
XX	
DE	S. mutans glucosyltransferase GTFB mutant K1014T.
XX	
KM	Glucosyltransferase; GTFB; transgenic plant; paper sizing;
KM	coating composition; glucan; starch; latex; thermoplastic molecule;
KW	amyloplast; vacuole; paper manufacture; mutant; mutcin.
XX	
OS	Streptococcus mutans.
OS	Synthetic.
XX	
FH	Key Location/Qualifiers
FT	Misc-difference 1014
FT	/note= "Wild-type Lys substituted by Thr"
XX	
PN	US2002031826-A1.
XX	
PD	14-MAR-2002.
XX	
PF	19-DEC-2000; 2000US-0740274.
XX	
PR	11-DEC-1998; 98US-0210361.
PR	07-JUN-1995; 95US-0478704.
PR	07-JUN-1995; 95US-0482711.
PR	07-JUN-1995; 95US-0485243.
PR	16-JAN-1998; 98US-0007959.
PR	16-JAN-1998; 98US-0008172.
PR	16-JAN-1998; 98US-0009620.
XX	
PA	(NICH/) NICHOLS S E.
XX	
PI	Nichols SE;
XX	
DR	WPI; 2002-414332/44.
XX	
PT	Glucosyltransferase B or D protein useful for producing a glucan useful
PT	as substitutes for and additions to modified starch and latexes in
PT	paper manufacture, comprises mutations in specific positions
XX	
PS	Claim 36; Page -; 44pp; English.
XX	
CC	The invention an isolated protein comprising a glucosyltransferase
CC	(GTF) B polypeptide having changes at position from 1448V, D457N,

[illegible]

XX 14-MAR-2002.
 PD 19-DEC-2000; 2000US-0740274.
 PF 11-DEC-1998; 98US-0210361.
 XX 07-JUN-1995; 95US-0478704.
 PR 07-JUN-1995; 95US-0482711.
 PR 07-JUN-1995; 95US-0485243.
 PR 16-JAN-1998; 98US-0007999.
 PR 16-JAN-1998; 98US-0008172.
 PR 20-JAN-1998; 98US-0009620.
 XX (NICH/) NICHOLS S E.
 PA Nichols SE;
 XX WPI; 2002-414332/44.
 XX Glucosyltransferase B or D protein useful for producing a glucan useful
 PT as substitutes for and additions to modified starch and latexes in
 PT paper manufacture, comprises mutations in specific positions -
 PS Claim 36; Page -; 44pp; English.
 XX The invention an isolated protein comprising a glucosyltransferase
 CC (GTF) B polypeptide having changes at position from 1448V, D457N,
 CC D567T, K1014T, D457N/D567T, D457N/D571K, D567T/D571K,
 CC D567T/D571K/K1014T, T448V/D457N/D567T/D571K/K790/K1014T,
 CC Y169A/Y170A/Y171A, and K779Q or a GTF D polypeptide having
 CC changes at positions from T589D, T589E, N471D, N471D/T589D, and
 CC N471D/T589E. Also included are a glucan produced by the GTF mutant,
 CC an isolated polynucleotide which encodes P1 or P2, or its complementary
 CC polynucleotide, a ribonucleic acid sequence encoding the GTF mutant,
 CC an expression cassette comprising the polynucleotide operably linked to a
 CC promoter, a vector comprising the expression cassette, host cell
 CC introduced with the vector, a transgenic plant comprising the
 CC vector, a seed or tuber from the transgenic plant, a paper sizing and/or
 CC coating composition comprising a glucan produced in a plant transformed
 CC with a gene encoding the mutant GTF, wild type or, starch, a latex,
 CC thermoplastic molecule or their combinations or glucan and starch where
 CC the glucan is produced in the amyloplast and/or vacuole or a maize line
 CC deficient in starch biosynthesis, transformed with a gene encoding a
 CC glucosyltransferase B or D enzyme, wild-type or mutant and a paper
 CC comprising the glucan (paper sizing/coating agent). The vector is useful
 CC for producing a glucan in a plant. The method comprises transforming a
 CC plant cell with the vector, growing the plant cell under plant growing
 CC conditions to produce a regenerated plant and inducing expression of the
 CC polynucleotide for a time sufficient to produce the glucan in the
 CC regenerated plant, where the vector contains a transit sequence from
 CC ribulose biphosphate carboxylase small subunit, waxy, brittle-1 and
 CC chlorophyll AB binding protein to produce a transgenic plant, and glucan
 CC is produced in the amyloplast of potato or the vacuole of sugar beet.
 CC Glucans are useful as substitutes for and additions to modified starch
 CC and latexes in paper manufacture. Unlike prior art techniques, paper
 CC require input materials that produce chemical effluents, paper
 CC manufacture utilizing the glucan produced by GTF, which utilizes
 CC biologically produced input materials, is more cost-effective and
 CC environmentally friendly. Moreover, glucans also exhibit thermoplastic
 CC properties and impart gloss to the paper during coating step.
 CC The present sequence represents a GTF mutant of the invention.
 CC Note: The present sequence is not shown in the specification but
 CC was created by the indexer using the GTFB sequence appearing as AAU98027
 CC and the information in claim 36.
 XX Sequence 1475 AA;
 SQ Query Match 90.8%; Score 108; DB 23; Length 1475;
 Best Local Similarity 85.7%; Pred. No. 2.7e-07;
 Matches 18; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
 QY 1 ANNHVSIVFAMSDNDPPLYLHD 21
 ||:||||:|||||

DB 481 ANNHVSIVFAMSDNDPPLYLHD 501
 RESULT 9
 AAU98035
 ID AAU98035 standard; Protein; 1475 AA.
 XX AAU98035;
 AC 27-AUG-2002 (first entry)
 DT 27-AUG-2002 (first entry)
 XX S. mutans glucosyltransferase GTFB mutant D457N/D571K.
 DE Glucosyltransferase; GTFB; transgenic plant; paper sizing;
 KM coating composition; glucan; starch; latex; thermoplastic molecule;
 KM amyloplast; vacuole; paper manufacture; mutant; mutcin.
 XX Streptococcus mutans.
 OS Synthetic.
 XX Key Location/Qualifiers
 FH Misc-difference 457 /note= "Wild-type Asp substituted by Asn"
 FT Misc-difference 571 /note= "Wild-type Asp substituted by Lys"
 FT Misc-difference 571 /note= "Wild-type Asp substituted by Lys"
 XX US2002031826-A1.
 PN 14-MAR-2002.
 PD 19-DEC-2000; 2000US-0740274.
 XX 11-DEC-1998; 98US-0210361.
 PR 07-JUN-1995; 95US-0478704.
 PR 07-JUN-1995; 95US-0482711.
 PR 07-JUN-1995; 95US-0485243.
 PR 16-JAN-1998; 98US-0007999.
 PR 16-JAN-1998; 98US-0008172.
 PR 20-JAN-1998; 98US-0009620.
 XX (NICH/) NICHOLS S E.
 PA Nichols SE;
 XX WPI; 2002-414332/44.
 XX Glucosyltransferase B or D protein useful for producing a glucan useful
 PT as substitutes for and additions to modified starch and latexes in
 PT paper manufacture, comprises mutations in specific positions -
 PS Claim 36; Page -; 44pp; English.
 XX The invention an isolated protein comprising a glucosyltransferase
 CC (GTF) B polypeptide having changes at position from 1448V, D457N,
 CC D567T, K1014T, D457N/D567T, D457N/D571K, D567T/D571K,
 CC D567T/D571K/K1014T, T448V/D457N/D567T/D571K/K790/K1014T,
 CC Y169A/Y170A/Y171A, and K779Q or a GTF D polypeptide having
 CC changes at positions from T589D, T589E, N471D, N471D/T589D, and
 CC N471D/T589E. Also included are a glucan produced by the GTF mutant,
 CC an isolated polynucleotide which encodes P1 or P2, or its complementary
 CC polynucleotide, a ribonucleic acid sequence encoding the GTF mutant,
 CC an expression cassette comprising the expression cassette, host cell
 CC introduced with the vector, a transgenic plant comprising the
 CC vector, a seed or tuber from the transgenic plant, a paper sizing and/or
 CC coating composition comprising a glucan produced in a plant transformed
 CC with a gene encoding the mutant GTF, wild type or, starch, a latex,
 CC thermoplastic molecule or their combinations or glucan and starch where
 CC the glucan is produced in the amyloplast and/or vacuole or a maize line
 CC deficient in starch biosynthesis, transformed with a gene encoding a
 CC glucosyltransferase B or D enzyme, wild-type or mutant and a paper
 CC comprising the glucan (paper sizing/coating agent). The vector is useful
 CC for producing a glucan in a plant. The method comprises transforming a

CC plant cell with the vector, growing the plant cell under plant growing
 CC conditions to produce a regenerated plant and inducing expression of the
 CC polynucleotide for a time sufficient to produce the glucan in the
 CC regenerated plant, where the vector contains a transit sequence from
 CC ribulose biphosphate carboxylase small subunit, waxy, brittle-1 and
 CC chlorophyll AB binding protein to produce a transgenic plant, and glucan
 CC is produced in the amyloplast of potato or the vacuole of sugar beet.
 CC Glucans are useful as substitutes for and additions to modified starch
 CC and latexes in paper manufacture. Unlike prior art techniques, which
 CC require input materials that produce chemical effluents, paper
 CC manufacture utilizing the glucan produced by GTF, which utilizes
 CC biologically produced input materials, is more cost-effective and
 CC environmentally friendly. Moreover, glucans also exhibit thermoplastic
 CC properties and impart gloss to the paper during coating step.
 CC The present sequence represents a GTFB mutant of the invention.
 CC Note: The present sequence is not shown in the specification but
 CC was created by the indexer using the GTFB sequence appearing as AAU98027
 CC and the information in claim 36.

CC Sequence 1475 AA;

SO Query Match 90.8%; Score 108; DB 23; Length 1475;
 Best Local Similarity 85.7%; Pred. No. 2.7e-07;
 Matches 18; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 ANNHVSIIVKMSDNDPYLHD 21
 DB 481 ANDHLSILEKMSDNDPYLHD 501

RESULT 10

AAU98036 standard; Protein; 1475 AA.

AC AAU98036;

DT 27-AUG-2002 (first entry)

DE S. mutans glucosyltransferase GTFB mutant D567T/D571K.

XX Glucosyltransferase; GTFB; transgenic plant; paper sizing;
 XX coating composition; glucan; starch; latex; thermoplastic molecule;
 XX amyloplast; vacuole; paper manufacture; mutant; mutain.

XX Streptococcus mutans.

OS Synthetic.

PH Key Location/Qualifiers

FT Misc-difference 567 /note= "Wild-type Asp substituted by Thr"

FT Misc-difference 571 /note= "Wild-type Asp substituted by Lys"

FT US2002031826-A1.

XX 14-MAR-2002.

PD 19-DEC-2000; 2000US-0740274.

PR 11-DEC-1998; 98US-0210361.

PR 07-JUN-1995; 95US-0478704.

PR 07-JUN-1995; 95US-0482711.

PR 07-JUN-1995; 95US-0485243.

PR 16-JAN-1998; 98US-0007999.

PR 16-JAN-1998; 98US-0008172.

PR 20-JAN-1998; 98US-0009620.

XX (NICH/) NICHOLS S E.
 XX Nichols SE;
 FT
 DR WPI; 2002-414332/44.
 XX

PT Glucosyltransferase B or D protein useful for producing a glucan useful
 PT as substitutes for and additions to modified starch and latexes in
 PT paper manufacture, comprises mutations in specific positions
 PS Claim 36; Page -; 44pp; English.

XX The invention an isolated protein comprising a glucosyltransferase
 CC (GTF) B polypeptide having changes at position from 1448V, D457N,
 CC D567T, K1014T, D457N/D567T, D457N/D571K, D567T/D571K,
 CC D567T/K1014T, 1448V/D457N/D567T/D571K/K779Q/K1014T,
 CC Y163A/Y170K/Y171A, and K779Q or a GTF D polypeptide having
 CC changes at positions from T589D, T589E, N477D, N477D/T589D, and
 CC N477D/T589E. Also included are a glucan produced by the GTF mutant,
 CC an isolated polynucleotide which encodes P1 or P2, or its complementary
 CC polynucleotide, a ribonucleic acid sequence encoding the GTF mutant,
 CC an expression cassette comprising the polynucleotide operably linked to a
 CC promoter, a vector comprising the expression cassette, host cell
 CC introduced with the vector, a transgenic plant comprising the
 CC vector, a seed or tuber from the transgenic plant, a paper sizing and/or
 CC coating composition comprising a glucan produced in a plant transformed
 CC with a gene encoding the mutant GTF, wild type or, starch, a latex,
 CC thermoplastic molecule or their combinations or glucan and starch where
 CC the glucan is produced in the amyloplast and/or vacuole or a maize line
 CC deficient in starch biosynthesis, transformed with a gene encoding a
 CC glucosyltransferase B or D enzyme, wild-type or mutant and a paper
 CC comprising the glucan (paper sizing/coating agent). The vector is useful
 CC for producing a glucan in a plant. The method comprises transforming a
 CC plant cell with the vector, growing the plant cell under plant growing
 CC conditions to produce a regenerated plant and inducing expression of the
 CC polynucleotide for a time sufficient to produce the glucan in the
 CC regenerated plant, where the vector contains a transit sequence from
 CC ribulose biphosphate carboxylase small subunit, waxy, brittle-1 and
 CC chlorophyll AB binding protein to produce a transgenic plant, and glucan
 CC is produced in the amyloplast of potato or the vacuole of sugar beet.
 CC Glucans are useful as substitutes for and additions to modified starch
 CC and latexes in paper manufacture. Unlike prior art techniques, which
 CC require input materials that produce chemical effluents, paper
 CC manufacture utilizing the glucan produced by GTF, which utilizes
 CC biologically produced input materials, is more cost-effective and
 CC environmentally friendly. Moreover, glucans also exhibit thermoplastic
 CC properties and impart gloss to the paper during coating step.
 CC The present sequence represents a GTFB mutant of the invention.
 CC Note: The present sequence is not shown in the specification but
 CC was created by the indexer using the GTFB sequence appearing as AAU98027
 CC and the information in claim 36.

SO Sequence 1475 AA;

Query Match 90.8%; Score 108; DB 23; Length 1475;
 Best Local Similarity 85.7%; Pred. No. 2.7e-07;
 Matches 18; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 ANNHVSIIVKMSDNDPYLHD 21
 DB 481 ANDHLSILEKMSDNDPYLHD 501

RESULT 11

AAU98037 standard; Protein; 1475 AA.

AC AAU98037;

DT 27-AUG-2002 (first entry)

DE S. mutans glucosyltransferase GTFB mutant D567T/D571K/K1014T.

XX Glucosyltransferase; GTFB; transgenic plant; paper sizing;
 XX coating composition; glucan; starch; latex; thermoplastic molecule;
 XX amyloplast; vacuole; paper manufacture; mutant; mutain.

XX Streptococcus mutans.
 OS Synthetic.

XX Key Location/Qualifiers
 FH Misc-difference 567 /note= "Wild-type Asp substituted by Thr"
 FT Misc-difference 571 /note= "Wild-type Asp substituted by Lys"
 FT Misc-difference 1014 /note= "Wild-type Lys substituted by Thr"
 XX US2002031826-A1.
 XX 14-MAR-2002.
 XX 19-DEC-2000; 2000US-0740274.
 XX 11-DEC-1998; 98US-0210361.
 XX 07-JUN-1995; 95US-0478704.
 XX 07-JUN-1995; 95US-0482711.
 XX 07-JUN-1995; 95US-0485243.
 XX 16-JAN-1998; 98US-0007999.
 XX 16-JAN-1998; 98US-0008172.
 XX 20-JAN-1998; 98US-0009620.
 XX (NICH/) NICHOLS S E.
 XX Nichols SE;
 XX WPI: 2002-414332/44.
 XX Glucosyltransferase B or D protein useful for producing a glucan useful
 PT as substitutes for and additions to modified starch and latexes in
 PT paper manufacture, comprises mutations in specific positions -
 XX
 PS Claim 36; Page -; 44pp; English.
 XX The invention an isolated protein comprising a glucosyltransferase
 CC (GTF) B polypeptide having changes at position from I448V, D457N,
 CC D567T, K1014T, D457N/D567T, D457N/D571K, D567T/D571K,
 CC D567T/D571K/K1014T, I448V/D457N/D567T/K779Q/K1014T,
 CC Y169N/Y170N/Y171A, and K779Q or a GTF D polypeptide having
 CC changes at positions from T589D, T589E, N471D, N471D/T589D, and
 CC N471D/T589E. Also included are a glucan produced by the GTF mutant,
 CC an isolated polynucleotide which encodes P1 or P2, or its complementary
 CC polynucleotide, a ribonucleic acid sequence encoding the GTF mutant
 CC an expression cassette comprising the polynucleotide operably linked to a
 CC promoter, a vector comprising the expression cassette, host cell
 CC introduced with the vector, a transgenic plant comprising the
 CC vector, a seed or tuber from the transgenic plant, a paper sizing and/or
 CC coating composition comprising a glucan produced in a plant transformed
 CC with a gene encoding the mutant GTF, wild type or, starch, a latex, where
 CC the thermoplastic molecule or their combinations or glucan and starch where
 CC the glucan is produced in the amyloplast and/or vacuole or a maize like
 CC deficient in starch biosynthesis, transformed with a gene encoding a
 CC glucosyltransferase B or D enzyme, wild-type or mutant and a paper
 CC comprising the glucan (paper sizing/coating agent). The vector is useful
 CC for producing a glucan in a plant. The method comprises transforming a
 CC plant cell with the vector, growing the plant cell under plant growing
 CC conditions to produce a regenerated plant and inducing expression of the
 CC polynucleotide for a time sufficient to produce the glucan in the
 CC regenerated plant, where the vector contains a transit sequence from
 CC ribulose biphosphate carboxylase small subunit, waxy, brittle-1 and
 CC chlorophyll AB binding protein to produce a transgenic plant, and glucan
 CC is produced in the amyloplast of potato or the vacuole of sugar beet.
 CC Glucans are useful as substitutes for and additions to modified starch
 CC and latexes in paper manufacture. Unlike prior art techniques, which
 CC require input materials that produce chemical effluents, paper
 CC manufacture utilizing the glucan produced by GTF, which utilizes
 CC biologically produced input materials, is more cost-effective and
 CC environmentally friendly. Moreover, glucans also exhibit thermoplastic
 CC properties and impart gloss to the paper during coating step.
 CC The present sequence represents a GTFB mutant of the invention.
 CC Note: The present sequence is not shown in the specification but
 CC was created by the indexer using the GTFB sequence appearing as AAU96027

CC and the information in claim 36.
 XX SQ Sequence 1475 AA;
 XX Query Match 90.8%; Score 108; DB 23; Length 1475;
 XX Best Local Similarity 85.7%; Pred. No. 2.7e-07;
 XX Matches 18; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
 XX QY 1 ANNHVSIVKASNDPRLYLD 21
 XX DB 481 ANDHSLIKWASNDPRLYLD 501
 XX
 XX RESULT 12
 XX ID AAU98038 standard; Protein; 1475 AA.
 XX AC AAU98038;
 XX DT 27-AUG-2002 (first entry)
 XX DE S. mutans GTFB mutant I448V/D457N/D567T/D571K/K779Q/K1014T.
 XX KW Glucosyltransferase; GTFB; transgenic plant; paper sizing;
 XX KM coating composition; glucan; starch; latex; thermoplastic molecule;
 XX amyloplast; vacuole; paper manufacture; mutant; mutcin.
 XX OS Streptococcus mutans.
 XX OS Synthetic.
 XX
 XX Key Location/Qualifiers
 FH Misc-difference 448 /note= "Wild-type Ile substituted by Val"
 FT Misc-difference 457 /note= "Wild-type Ile substituted by Asn"
 FT Misc-difference 567 /note= "Wild-type Asp substituted by Asn"
 FT Misc-difference 571 /note= "Wild-type Asp substituted by Thr"
 FT Misc-difference /note= "Wild-type Asp substituted by Lys"
 FT Misc-difference 779 /note= "Wild-type Lys substituted by Gln"
 FT Misc-difference 1014 /note= "Wild-type Lys substituted by Thr"
 FT
 FT
 XX US2002031826-A1.
 XX 14-MAR-2002.
 XX 19-DEC-2000; 2000US-0740274.
 XX 11-DEC-1998; 98US-0210361.
 XX 07-JUN-1995; 95US-0478704.
 XX 07-JUN-1995; 95US-0482711.
 XX 07-JUN-1995; 95US-0485243.
 XX 16-JAN-1998; 98US-0007999.
 XX 16-JAN-1998; 98US-0008172.
 XX 20-JAN-1998; 98US-0009620.
 XX
 XX (NICH/) NICHOLS S E.
 XX Nichols SE;
 XX WPI: 2002-414332/44.
 XX Glucosyltransferase B or D protein useful for producing a glucan useful
 PT as substitutes for and additions to modified starch and latexes in
 PT paper manufacture, comprises mutations in specific positions -
 XX
 PS Claim 36; Page -; 44pp; English.
 XX The invention an isolated protein comprising a glucosyltransferase
 CC (GTF) B polypeptide having changes at position from I448V, D457N,
 CC D567T, K1014T, D457N/D567T, D457N/D571K, D567T/D571K,

RESULT 14
AAU98040
AAU98040 standard; Protein; 1475 AA.
XX
AC AAU98040;
XX
XX 27-AUG-2002 (first entry)
DE
XX S. mutans glucosyltransferase GTFB mutant K779Q.
XX
XX Glucosyltransferase; GTFB; transgenic plant; paper sizing;
XX coating composition; glucan; starch; latex; thermoplastic molecule;
XX amyloplast; vacuole; paper manufacture; mutant; mutein.
XX
XX Streptococcus mutans.
OS Synthetic.
XX
XX Key Location/Qualifiers
XX MISC-difference 779 /note= "wild-type Lys substituted by Gln"
XX
XX US2002031826-A1.
XX
XX 14-MAR-2002.
XX
XX 19-DEC-2000; 2000US-0740274.
XX
XX 11-DEC-1998; 98US-0210361.
XX 07-JUN-1995; 95US-0478704.
XX 07-JUN-1995; 95US-0482711.
XX 07-JUN-1995; 95US-0485243.
XX 16-JAN-1998; 98US-0007999.
XX 16-JAN-1998; 98US-0008172.
XX 20-JAN-1998; 98US-0009620.
XX
XX (NICH/) NICHOLS S E.
XX
XX Nichols SE;
XX
XX WPI; 2002-414332/44.
XX
XX Glucosyltransferase B or D protein useful for producing a glucan useful
PT as substrates for and additions to modified starch and latexes in
PT paper manufacture, comprises mutations in specific positions -
XX
XX Claim 36; Page -; 44pp; English.
XX
XX The invention an isolated protein comprising a glucosyltransferase
CC (GTF) B polypeptide having changes at position from 1448V, D457N,
CC D567T, K1014T, D457N/D567T, D457N/D571K, D567T/D571K,
CC D567T/D571K/K1014T, 1448V/D457N/D567T/D571K/K779Q/K1014T,
CC Y169A/Y170A/Y171A, and K779Q or a GTF D polypeptide having
CC changes at positions from T589D, T589E, N471D, N471D/T589D, and
CC N471D/T589E. Also included are a glucan produced by the GTF mutant,
CC an isolated polynucleotide which encodes P1 or P2, or its complementary
CC polynucleotide, a ribonucleic acid sequence encoding the GTF mutant,
CC an expression cassette comprising the polynucleotide operably linked to a
CC promoter, a vector comprising the expression cassette, host cell
CC introduced with the vector, a transgenic plant comprising the
CC vector, a seed or tuber from the transgenic plant, a paper sizing and/or
CC coating composition comprising a glucan produced in a plant transformed
CC with a gene encoding the mutant GTF wild type or, starch, a latex,
CC thermoplastic molecule or their combinations or glucan and starch where
CC the glucan is produced in the amyloplast and/or vacuole or a maize line
CC deficient in starch biosynthesis, transformed with a gene encoding a
CC glucosyltransferase B or D enzyme, wild-type or mutant and a paper
CC comprising the glucan (paper sizing/coating agent). The vector is useful
CC for producing a glucan in a plant. The method comprises transforming a
CC plant cell with the vector, growing the plant cell under plant growing
CC conditions to produce a regenerated plant and inducing expression of the
CC polynucleotide for a time sufficient to produce the glucan in the
CC regenerated plant, where the vector contains a transit sequence from

CC ribulose biphosphate carboxylase small subunit, waxy, brittle-1 and
CC chlorophyll AB binding protein to produce a transgenic plant, and glucan
CC is produced in the amyloplast of potato or the vacuole of sugar beet.
CC Glucans are useful as substitutes for and additions to modified starch
CC and latexes in paper manufacture. Unlike prior art techniques, which
CC require input materials that produce chemical effluents, paper
CC manufacture utilizing the glucan produced by GTF, which utilizes
CC biologically produced input materials, is more cost-effective and
CC environmentally friendly. Moreover, glucans also exhibit thermoplastic
CC properties and impart gloss to the paper during coating step.
CC The present sequence represents a GTFB mutant of the invention.
CC Note: The present sequence is not shown in the specification but
CC was created by the indexer using the GTFB sequence appearing as AAU98027
CC and the information in Claim 36.
XX
XX
XX Sequence 1475 AA;
SQ
Query March 90.8%; Score 108; DB 23; Length 1475;
Best Local Similarity 85.7%; Pred. No. 2.7e-07;
Matches 18; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
OY 1 ANNHVSIVEAMSDNDTPYLHD 21
DB 481 ANNHVSIVEAMSDNDTPYLHD 501
RESULT 15
AAU9284
ID AAU9284 standard; Protein; 1476 AA.
XX
XX AAU9284;
AC
XX 13-AUG-2002 (first entry)
XX
XX Streptococcus mutans monoclonal antibody-related protein #1.
DE
XX Streptococcus mutans monoclonal antibody-related protein #1.
XX Antibody; dental caries; water insoluble glucan synthetase;
XX anti-caries; glucosyl transferase-B; immunotherapy.
XX
XX Streptococcus mutans.
OS
XX JP2002114709-A.
XX
XX 16-APR-2002.
XX
XX 04-OCT-2000; 2000JP-0304889.
XX
XX 04-OCT-2000; 2000JP-0304889.
XX
XX (UYN1-) UNITV NIPPON.
XX
XX WPI; 2002-448101/48.
XX
XX Anti-carries agent composed of a monoclonal antibody against an
PT inhibitory enzyme against water insoluble glucan synthetase of glucosyl
PT transferase-B (GTF-B) of Streptococcus mutans -
XX
XX Claim 3; Page 13-16; 28pp; Japanese.
XX
XX The invention relates to a monoclonal antibody against dental caries and
CC an anti-carries agent composed of a monoclonal antibody produced by
CC Streptococcus mutans, particularly mouse-hybridoma MHP126 (FERM P-17566)
CC or mouse-hybridoma MHP136 (FERM P-17567), against an enzyme having
CC inhibitive activity against water insoluble glucan synthetase of glucosyl
CC transferase-B. The monoclonal antibody specifically inhibits water
CC insoluble glucan synthetase of Streptococcus mutans produced glucosyl
CC transferase-B and is used in the immunotherapy of dental caries. This
CC sequence represents a Streptococcus mutans monoclonal antibody-related
CC protein.
XX
XX Sequence 1476 AA;
SQ
Query March 90.8%; Score 108; DB 23; Length 1476;

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 13, 2003, 09:31:40 ; Search time 13.0379 Seconds
(without alignments)
154.898 Million cell updates/sec

Title: US-09-290-049a-12
Perfect score: 119
Sequence: 1 ANNHVSIVEAWSNDNTPYLHD 21

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :
1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	119	100.0	A38175	glucosyltransferase
2	108	90.8	A38175	glfB protein precursor
3	99	83.2	JT0345	dextranucrase (EC
4	74	62.2	A45866	probable dextran
5	68	57.1	T31098	glucosyltransferase
6	67	56.3	A41483	glucosyltransferase
7	61	51.3	T30857	glucosyltransferase
8	61	51.3	T30552	glucosyltransferase
9	57	47.9	T30858	glucosyltransferase
10	56	47.1	A44811	glucosyltransferase
11	55	46.2	S22737	hypothetical prote
12	55	44.5	C86205	BglF2 protein - nu
13	49	41.2	Q08E40	conserved hypotet
14	49	41.2	B82220	glycerinaldehyde-3-p
15	48	40.3	A38175	hypothetical prote
16	46	38.7	S59448	probable pdk prot
17	46	38.7	H70538	pyruvate, phosphat
18	46	38.7	E87028	4-aminobutyrate am
19	45.5	38.2	D90506	hypothetical prote
20	45	37.8	C64233	RogG domain cont
21	45	37.8	T48610	hypothetical prote
22	45	37.8	S25	hypothetical prote
23	45	37.8	E86362	hypothetical prote
24	45	37.8	T02748	pristinamycin I sy
25	45	37.8	T30289	site-specific DNA-
26	44.5	37.4	835	metalloproteinase
27	44	37.0	A59414	X-pro dipeptidyl-p
28	44	37.0	UC5142	probable membrane
29	44	37.0	S64443	

30	44	37.0	1131	2	T16217	hypothetical prote
31	44	37.0	1313	2	T29027	hypothetical prote
32	43.5	36.6	491	1	D64947	glucose-6-phosphat
33	43.5	36.6	491	2	B90949	glucose-6-phosphat
34	43.5	36.6	491	2	F85797	glucose-6-phosphat
35	43.5	36.6	491	2	AB0742	hypothetical prote
36	43	36.1	78	2	S76593	conserved hypotet
37	43	36.1	194	2	H72037	hypothetical prote
38	43	36.1	194	2	C66586	hypothetical prote
39	43	36.1	275	2	S55978	heat-shock sigma f
40	43	36.1	284	2	JC7148	glycerinaldehyde-3-p
41	43	36.1	331	1	DEUTGC	hypothetical prote
42	43	36.1	346	2	E70715	hypothetical prote
43	43	36.1	351	2	D69991	probable glutathio
44	43	36.1	396	2	T35024	probable homogenit
45	43	36.1	441	2	T36737	

ALIGNMENTS

RESULT 1
A38175
glucosyltransferase precursor - Streptococcus sobrinus
C/Species: Streptococcus sobrinus
C/Date: 28-Aug-1992 #sequence_revision 28-Aug-1992 #text_change 15-Oct-1999
R/Accession: A38175
R/AbO: H.; Mtsunura, T.; Kodama, T.; Ohta, H.; Fukui, K.; Kato, K.; Kagawa, H.
U. Bacteriol. 173, 989-996, 1991
A/Title: Peptide sequences for sucrose splitting and glucan binding within Streptococcus
A/Reference number: A38175; MUID:91123227; PMID:1704006
A/Accession: A38175
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-1592 <ABO>
A/Cross-references: GB:D90213; NID:9217032; PIDN:BA414241.1; PID:dl014946; PID:9217033
C/Superfamily: cpl repeat homology
F/1093-1112/Domain: cpl repeat homology <CP1>
F/1222-1241/Domain: cpl repeat homology <CP2>
F/1287-1306/Domain: cpl repeat homology <CP3>
F/1330-1351/Domain: cpl repeat homology <CP4>
F/1352-1371/Domain: cpl repeat homology <CP5>
F/1402-1420/Domain: cpl repeat homology <CP6>
F/1465-1484/Domain: cpl repeat homology <CP7>
F/1513-1532/Domain: cpl repeat homology <CP8>

Query Match 100.0%; Score 119; DB 2; Length 1592;
Best Local Similarity 100.0%; Pred. No. 5.6e-10;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ANNHVSIVEAWSNDNTPYLHD 21
Db 477 ANNHVSIVEAWSNDNTPYLHD 497

RESULT 2
B33135
glfB protein precursor - Streptococcus mutans
C/Species: Streptococcus mutans
C/Date: 23-Oct-1990 #sequence_revision 23-Oct-1990 #text_change 15-Oct-1999
R/Accession: B33135
R/Shirota, T.; Ueda, S.; Kuramitsu, H.K.
U. Bacteriol. 169, 4263-4270, 1987
A/Title: Sequence analysis of the glfB gene from Streptococcus mutans.
A/Reference number: A33135; MUID:87308013; PMID:3040685
A/Accession: B33135
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-1475 <SHI>
A/Cross-references: GB:M17361; NID:g153639; PIDN:AAA86588.1; PID:g153640
R/Shirota, T.; Ueda, S.; Kuramitsu, H.K.
submitted to the Protein Sequence Database, September 1990
A/Reference number: A33128

A:Accession: A33128
 A:Status: preliminary; not compared with conceptual translation
 A:Molecule type: DNA
 A:Residues: 1-171,173-641,'N',643-1475 <SH2>
 A:Experimental source: strain GS-5
 C:Superfamily: cpl repeat homology
 F:1096-1115/Domain: cpl repeat homology <CP1>
 F:1124-1243/Domain: cpl repeat homology <CP2>
 F:1289-1308/Domain: cpl repeat homology <CP3>
 F:1354-1373/Domain: cpl repeat homology <CP4>
 F:1419-1438/Domain: cpl repeat homology <CP5>

Query Match 90.8%; Score 108; DB 2; Length 1475;
 Best Local Similarity 85.7%; Pred. No. 2.6e-08;
 Matches 18; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 ANNHVSIVEAMSDNDTPYLHD 21
 Db 481 ANDHLSITLSEAMSDNDTPYLHD 501

RESULT 3
 J03045 dextranucrase (EC 2.4.1.5) precursor - Streptococcus mutans (strain GS-5)
 N:Alternate names: sucrose 6-glycosyltransferase
 C:Species: Streptococcus mutans
 C>Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 24-Sep-1999
 C:Accession: J03045; C33135
 R:Ueda, S.; Shiroza, T.; Kuramitsu, H.K.
 Gene 69, 101-109, 1988
 A>Title: Sequence analysis of the glfC gene from Streptococcus mutans GS-5.
 A:Reference number: J03045; MUID:89137980; PMID:2976010
 A:Accession: J03045
 A:Molecule type: DNA
 A:Residues: 1-1375 <UED>
 A:Experimental source: GS-5
 R:Shiroza, T.; Ueda, S.; Kuramitsu, H.K.
 J. Bacteriol. 169, 4263-4270, 1987
 A>Title: Sequence analysis of the glfB gene from Streptococcus mutans.
 A:Reference number: A33135; MUID:87308013; PMID:3040685
 A:Accession: C33135
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-349 <SH1>
 A:Cross-references: GB:M17361
 C:Genetics:
 A:Gene: glfC
 C:Function:
 A:Description: catalyzes the synthesis of both water-soluble and water-insoluble glucans
 C:Superfamily: cpl repeat homology
 C:Keywords: duplication; glycosyltransferase; hexosyltransferase
 F:1-34/Domain: signal sequence #status predicted <Sig>
 F:35-1375/Product: glucosyltransferase #status predicted <Mat>
 F:1126-1145/Domain: cpl repeat homology <CP1>
 F:1253-1272/Domain: cpl repeat homology <CP2>
 F:1318-1337/Domain: cpl repeat homology <CP3>

Query Match 83.2%; Score 99; DB 2; Length 1375;
 Best Local Similarity 81.0%; Pred. No. 6.1e-07;
 Matches 17; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 ANNHVSIVEAMSDNDTPYLHD 21
 Db 507 ANDHLSITLSEAMSDNDTPYLHD 527

RESULT 4
 A43866 dextranucrase (EC 2.4.1.5) precursor - Streptococcus mutans
 C:Species: Streptococcus mutans
 C>Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 07-May-1999
 C:Accession: A43866
 R:Honda, O.; Kato, C.; Kuramitsu, H.K.

J. Gen. Microbiol. 136, 2099-2105, 1990
 A>Title: Nucleotide sequence of the Streptococcus mutans glfD gene encoding the glucosyl
 A:Reference number: A45866; MUID:91100958; PMID:2148600
 A:Accession: A45866
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-1431 <HON>
 A:Cross-references: GB:M29296
 C:Superfamily: cpl repeat homology
 C:Keywords: glycosyltransferase; hexosyltransferase
 F:181-201/Domain: cpl repeat homology <CP1>
 F:1127-1146/Domain: cpl repeat homology <CP2>
 F:1192-1211/Domain: cpl repeat homology <CP3>
 F:1257-1276/Domain: cpl repeat homology <CP4>
 F:1277-1297/Domain: cpl repeat homology <CP5>
 F:1321-1340/Domain: cpl repeat homology <CP6>
 F:1341-1361/Domain: cpl repeat homology <CP7>
 F:1385-1404/Domain: cpl repeat homology <CP7>

Query Match 62.2%; Score 74; DB 2; Length 1431;
 Best Local Similarity 66.7%; Pred. No. 0.005;
 Matches 14; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 1 ANNHVSIVEAMSDNDTPYLHD 21
 Db 495 ANHLSITLSEAMSDNDTPYLHD 515

RESULT 5
 T31098 probable dextranucrase (EC 2.4.1.5), extracellular - Leuconostoc mesenteroides
 C:Species: Leuconostoc mesenteroides
 C>Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 11-May-2000
 C:Accession: T31098
 R:Monchois, V.; Remaud-Simeon, M.; Monsar, P.; Willemot, R.M.
 FEMS Microbiol. Lett. 159, 307-315, 1998
 A>Title: Cloning and sequencing of a gene coding for an extracellular dextranucrase (DS
 A:Reference number: Z20981; MUID:98164374; PMID:9503626
 A:Accession: T31098
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-1508 <MON>
 A:Cross-references: EMBL:AF030129; NID:g2766611; PID:g2766612; PIDN:AA95453.1
 A:Experimental source: strain NRRL B-1299
 C:Genetics:
 A:Gene: dsrB
 C:Function:
 A:Description: produces dextran composed only of alpha(1-6) glucosidic bonds
 C:Keywords: glycosyltransferase; hexosyltransferase

Query Match 57.1%; Score 68; DB 2; Length 1508;
 Best Local Similarity 57.1%; Pred. No. 0.046;
 Matches 12; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 1 ANNHVSIVEAMSDNDTPYLHD 21
 Db 563 ANOHLSTLEDMSHNDPEYKXD 583

RESULT 6
 A41483 glucosyltransferase (EC 2.4.1.-) glfS precursor - Streptococcus sobrinus
 C:Species: Streptococcus sobrinus
 C>Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 15-Oct-1999
 C:Accession: A41483
 R:Gilmore, K.S.; Russell, R.R.B.; Ferretti, J.J.
 Infect. Immun. 58, 2452-2458, 1990
 A>Title: Analysis of the Streptococcus downei glfS gene, which specifies a glucosyltran-
 A:Reference number: A41483; MUID:90316665; PMID:2142479
 A:Accession: A41483
 A:Molecule type: DNA
 A:Residues: 1-1365 <GLI>
 A:Cross-references: GB:M30943; NID:g153652; PIDN:AAA6898.1; PID:g153653

C:Genetics:
A:Gene: gtfS
C:Superfamily: cpl repeat homology
C:Keywords: glycosyltransferase; hexosyltransferase

Query Match 56.3%; Score 67; DB 2; Length 1365;
Best Local Similarity 57.1%; Pred. No. 0.058;
Matches 12; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 1 ANNHVSIVEAMSDNDTPYLHD 21
Db 467 AIDHLSTLEAMSGNDNDYKMD 487

RESULT 7
T30857
glucosyltransferase - Streptococcus salivarius

C:Species: Streptococcus salivarius

C:Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 22-Oct-1999

C:Accession: T30857

R:Simpson, C.L.; Giffard, P.M.; Jacques, N.A.

Infected. Immun. 63, 609-621, 1995

A:Title: Streptococcus salivarius ATCC 25975 possesses at least two genes coding for pri

A:Reference number: Z20909; MUID:95122197; PMID:7822030

A:Accession: T30857

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-1449 <SIM>

A:Cross-references: EMBL:LJ35495; NID:G662378; PID:G662379; PIDN:AAC41412.1

A:Gene: gtfL

Query Match 51.3%; Score 61; DB 2; Length 1449;
Best Local Similarity 57.1%; Pred. No. 0.54;
Matches 12; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 1 ANNHVSIVEAMSDNDTPYLHD 21
Db 536 AIKHLSTLEAMSGNDNDYKMD 556

RESULT 8
T30552
glucosyltransferase N - Streptococcus salivarius (fragment)

C:Species: Streptococcus salivarius

C:Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 22-Oct-1999

C:Accession: T30552

R:Jaffe, R.I.

submitted to the EMBL Data Library, February 1998

A:Description: Streptococcus salivarius V1477 gtfN.

A:Reference number: Z20854

A:Accession: T30552

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-1449 <JAF>

A:Cross-references: EMBL:AF049609; NID:G2935545; PID:G2935546; PIDN:AAC05156.1

C:Genetics:
A:Gene: gtfN

Query Match 51.3%; Score 61; DB 2; Length 1449;
Best Local Similarity 57.1%; Pred. No. 0.54;
Matches 12; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 1 ANNHVSIVEAMSDNDTPYLHD 21
Db 536 AIKHLSTLEAMSGNDNDYKMD 556

RESULT 9
T30858
glucosyltransferase - Streptococcus salivarius

C:Species: Streptococcus salivarius

C:Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 22-Oct-1999

C:Accession: T30858
R:Simpson, C.L.; Giffard, P.M.; Jacques, N.A.

Infected. Immun. 63, 609-621, 1995

A:Title: Streptococcus salivarius ATCC 25975 possesses at least two genes coding for pri

A:Reference number: Z20909; MUID:95122197; PMID:7822030

A:Accession: T30858

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-1577 <SIM>

A:Cross-references: EMBL:LJ3528; NID:G662380; PID:G662381; PIDN:AAC41413.1

A:Gene: gtfM

Query Match 47.9%; Score 57; DB 2; Length 1577;
Best Local Similarity 61.1%; Pred. No. 2.5;
Matches 11; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 4 HVSIVAMSDNDTPYLHD 21
Db 591 HSLTLEAMSYNDHQYKMD 608

RESULT 10
A44811
glucosyltransferase (EC 2.4.1.-) I - Streptococcus salivarius

C:Species: Streptococcus salivarius

C:Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 15-Oct-1999

C:Accession: A44811; S22726; S28809

R:Giffard, P.M.; Simpson, C.L.; Milward, C.P.; Jacques, N.A.

J. Gen. Microbiol. 137, 2577-2593, 1991

A:Title: Molecular characterization of a cluster of at least two glucosyltransferase gen

A:Reference number: A44811; MUID:92148377; PMID:1838391

A:Accession: A44811

A:Molecule type: DNA

A:Residues: 1-1518 <GIF>

A:Cross-references: EMBL:Z11873; NID:G47526; PIDN:CAA77900.1; PID:G47527

A:Note: Sequence extracted from NCBI backbone (NCBI:G1050, NCBI:P:G1052)

C:Genetics:
A:Gene: gtfL

C:Superfamily: cpl repeat homology

C:Keywords: glycosyltransferase; hexosyltransferase

F:1307-1326/Domain: cpl repeat homology <CP4>

Query Match 47.1%; Score 56; DB 2; Length 1518;
Best Local Similarity 60.0%; Pred. No. 3.4;
Matches 9; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 4 HVSIVAMSDNDTPYLHD 18
Db 504 HVSIVAMSLNDNHV 518

RESULT 11
S22737
glucosyltransferase (EC 2.4.1.-) S - Streptococcus salivarius

C:Species: Streptococcus salivarius

C:Date: 16-Sep-1992 #sequence_revision 16-Sep-1992 #text_change 21-Jan-2000

C:Accession: S22737; S28810; B44811; S22727

R:Jacques, N.

submitted to the EMBL Data Library, March 1992

A:Reference number: S22726

A:Accession: S22737

A:Molecule type: DNA

A:Residues: 1-1599 <JAC>

A:Cross-references: EMBL:Z11872; NID:G47530; PIDN:CAA77898.1; PID:G47531

A:Experimental source: ATCC 25975

R:Giffard, P.M.; Simpson, C.L.; Milward, C.P.; Jacques, N.A.

J. Gen. Microbiol. 137, 2577-2593, 1991

A:Title: Molecular characterization of a cluster of at least two glucosyltransferase gen

A:Reference number: A44811; MUID:92148377; PMID:1838391

A:Accession: S28810

A:Molecule type: DNA

A:Residues: 1-51 <GIF>

A:Cross-references: EMBL:Z11873
A:Genetics: gtrk
A:Gene: gtrk
C:Superfamily: cpl repeat homology
C:Keywords: glycosyltransferase; hexosyltransferase
F:1456-1475/Domain: cpl repeat homology <CPR>

Query Match 46.2%; Score 55; DB 2; Length 1599;
Best Local Similarity 50.0%; Pred. No. 5.2;
Matches 9; Conservative 6; Mismatches 3; Indels 0; Gaps 0;

QY 4 HVSIVAEWSDNDTPYLHD 21
DB 494 NISLIEAWSHNDPIYVNE 511

RESULT 12
C66205
hypothetical protein [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
C:Accession: C66205
R:Theologian, A.; Becker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.E.; Dewar, K.;
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000

A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Maiti, R.; Marziani,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzbach, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MUID:21016719; PMID:11130712
A:Accession: C66205

A:Status: preliminary
A:Molecule type: DNA

A:Residues: 1-175 <STO>

A:Cross-references: GB:AE005172; NID:g9954041; PIDN:AAF62215.1; GSPDB:GN00.41

C:Genetics:

A:Map position: 1

Query Match 44.5%; Score 53; DB 2; Length 175;
Best Local Similarity 42.1%; Pred. No. 0.78;
Matches 8; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

QY 3 NHVSIVAEWSDNDTPYLHD 21
DB 109 NHQEVIDAMSDHOKPLMTD 127

RESULT 13

Q08B40
BgLR2 protein - human herpesvirus 4 (strain B95-8)

C:Species: human herpesvirus 4; Epstein-Barr virus
C:Date: 25-Feb-1985 #sequence_revision 25-Feb-1985 #text_change 16-Jul-1999
C:Accession: C43044; J01381; A03784; A03794; S33036

R:Bankier, A.T.; Deininger, P.L.; Farrell, P.J.; Barrett, B.G.
Mol. Biol. Med. 1, 21-45, 1983

A:Title: Sequence analysis of the 17,166 bp EcoRI fragment C of B95-8 Epstein-Barr virus

A:Reference number: A93065; MUID:85035713; PMID:6092825

A:Accession: C43044

A:Molecule type: DNA

A:Residues: 1-336 <BAN>

A:Cross-references: EMBL:V01555; NID:g59074; PIDN:CAA24831.1; PID:g1334895

R:Bankier, A.T.; Biggin, M.D.; Deininger, P.L.; Farrell, P.J.; Gibson, T.J.; H
Nature 310, 207-211, 1984

A:Title: DNA sequence and expression of the B95-8 Epstein-Barr virus genome.

A:Reference number: A03794; MUID:84270667; PMID:6087149

A:Contents: annotation; protein coding region

A:Note: neither amino acid nor nucleotide sequence is given

R:Chen, M.R.; Hsu, T.Y.; Lin, S.W.; Chen, J.Y.; Yang, C.S.
J. Gen. Virol. 72, 3047-3055, 1991

A:Title: Cloning and characterization of cDNA clones corresponding to transcripts from B

A:Reference number: J01381; MUID:92113548; PMID:1662696

A:Accession: J01381

A:Molecule type: mRNA

A:Residues: 1-336 <CHE>

A:Cross-references: GB:S77132; NID:g243314; PIDN:AA21113.1; PID:g243315

C:Superfamily: Epstein-Barr virus BGLF2 protein

Query Match 41.2%; Score 49; DB 1; Length 336;
Best Local Similarity 41.2%; Pred. No. 7.1;
Matches 7; Conservative 6; Mismatches 4; Indels 0; Gaps 0;

QY 1 ANNHVSIVAEWSDNDTP 17
DB 216 AGAHVNILRGMTEDDSP 232

RESULT 14
D82220
conserved hypothetical protein VCI268 [imported] - Vibrio cholerae (strain N16961 serog

C:Species: Vibrio cholerae
C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Feb-2001
C:Accession: D82220

R:Heidelberg, J.F.; Bisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwin, M.L.; Dodson, R.J.;
charlson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragol, I.; Sellers, F.
1, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
Nature 406, 477-483, 2000

A:Title: DNA sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
A:Reference number: A82035; MUID:20406833; PMID:10952301
A:Accession: D82220

A:Status: preliminary
A:Molecule type: DNA

A:Residues: 1-524 <HEI>

A:Cross-references: GB:AE004206; GB:AE003852; NID:g9655749; PIDN:AAF94427.1; GSPDB:GN001

A:Experimental source: serogroup O1; strain N16961; biotype El Tor

C:Genetics:

A:Gene: VCI268

A:Map position: 1

Query Match 41.2%; Score 49; DB 2; Length 524;
Best Local Similarity 43.5%; Pred. No. 12;
Matches 10; Conservative 4; Mismatches 7; Indels 2; Gaps 1;

QY 1 ANNHVSIVAEWSDNDTP 21
DB 365 ANNHWEILRGMSDPEVIDPALD 387

RESULT 15

B48445
glyceraldehyde-3-phosphate dehydrogenase (phosphorylating) (EC 1.2.1.12) - Leishmania me

C:Species: Leishmania mexicana
C:Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 03-Jun-2002
C:Accession: B48445; S25142

R:Hanaaert, V.; Blaauw, M.; Kohl, L.; Allert, S.; Opperdoes, F.R.; Michels, P.A.M.
Mol. Biochem. Parasitol. 55, 115-126, 1992

A:Title: Molecular analysis of the cytosolic and glycosomal glyceraldehyde-3-phosphate d

A:Reference number: A48445; MUID:93063042; PMID:1435864

A:Accession: B48445

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-331 <BAN>

A:Cross-references: EMBL:X65220; NID:g9552; PIDN:CAA46323.1; PID:g9553

C:Superfamily: glyceraldehyde-3-phosphate dehydrogenase

C:Keywords: oxidoreductase

Query Match 40.3%; Score 48; DB 2; Length 331;
Best Local Similarity 42.1%; Pred. No. 9.9;
Matches 8; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

QY 2 NNHVSIVAEWSDNDTPYLH 20
DB 301 NDHFVKLVSWYDNETGSH 319

Thu Nov 13 12:12:08 2003

us-09-290-049a-12.rpr

Page 5

Search completed: November 13, 2003, 09:50:26
Job time : 14.0379 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 13, 2003, 09:31:40 ; Search time 7.56398 Seconds
(without alignments)
130.561 Million cell updates/sec

Title: US-09-290-049a-12

Perfect score: 119
Sequence: 1 ANNHVSIYKASNDTPYLHD 21

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database: SwissProt_41:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	119	100.0	1 GF2_STRDO	P27470 streptococ
2	119	100.0	1 GF1_STRDO	P11001 streptococ
3	108	90.8	1 GF8_STRMU	P08987 streptococ
4	99	83.2	1 GF3C_STRMU	P13470 streptococ
5	74	62.2	1 GF2D_STRMU	P49331 streptococ
6	67	56.3	1 GF5_STRDO	P23336 streptococ
7	49	41.2	1 UL16_EBV	P03221 Epstein-Bar
8	48	40.3	1 GF3C_LEIME	Q01558 leishmania
9	46	38.7	1 YMS8_YEAST	Q03695 saccharomyc
10	45	37.8	1 GF3P_MYCGR	P47543 mycoplasma
11	44	37.0	1 SLIF_MOUSE	Q92093 mus musculu
12	44	37.0	1 YG3C_YEAST	P52280 saccharomyc
13	44	37.0	1 YPO4_CAEEL	P53015 caenorhadi
14	44	37.0	1 PCV_MOUSE	Q62059 mus musculu
15	43.5	36.6	1 G6PD_ECOLI	P22992 escherichia
16	43	36.1	1 SC35_YEAST	P53271 saccharomyc
17	43	36.1	1 GF3C_TRYRB	P10097 trypanosoma
18	43	36.1	1 Y943_MYCTU	P71566 mycobacteri
19	43	36.1	1 HBD_STRCO	Q92285 streptomyc
20	43	36.1	1 DOPO_HUMAN	P09172 homo sapien
21	43	36.1	1 SUHM_DROYI	Q08876 drosophila
22	42.5	35.7	1 AMP2_STRCO	Q60394 streptomyc
23	42.5	35.7	1 G6PD_ERMCH	P37986 erwinia chr
24	42	35.3	1 P8F1_DROME	Q92637 drosophila
25	42	35.3	1 CG16_YEAST	P22867 saccharomyc
26	42	35.3	1 GF3P_BACST	P00362 bacillus st
27	42	35.3	1 GF3P_BACCO	P11115 bacillus co
28	42	35.3	1 HGD_CAUCR	Q95588 caulobacter
29	42	35.3	1 HGD_PSEAR	Q95490 pseudomonas
30	42	35.3	1 GARD_PYRUF	O80400 pyrococcus
31	42	35.3	1 SYM_UREPA	Q92906 ureaplasma
32	42	35.3	1 YE09_MYCPN	P75375 mycoplasma
33	42	35.3	1 AD28_HUMAN	Q92637 homo sapien

34	42	35.3	776	1	AD28_MACFA	Q92637 macaca fasc
35	42	35.3	1627	1	ADP1_MYCPN	P11311 mycoplasma
36	42	35.3	3898	1	POIG_HCBV	P21530 hog cholera
37	41.5	34.9	416	1	HGD_LEGPN	Q95440 legionella
38	41.5	34.9	468	1	KG3H_DICDI	P51136 dictyostell
39	41.5	34.9	578	1	VAC8_YEAST	P39568 saccharomyc
40	41.5	34.9	684	1	ISH1_SCHPO	Q92746 schizosacch
41	41	34.5	130	1	UCR6_SCHMA	O01374 schistosoma
42	41	34.5	236	1	UCR4_TOBAC	P51134 nicotiana t
43	41	34.5	241	1	HEP1_HAEIN	P35757 haemophilus
44	41	34.5	241	1	HEP2_HAEIN	P45991 haemophilus
45	41	34.5	301	1	Y186_MYCPN	P75265 mycoplasma

ALIGNMENTS

```

RESULT 1
GF2_STRDO          STANDARD;          PRT; 1592 AA.
ID  GF2_STRDO
AC  P27470;
DT  01-AUG-1992 (Rel. 23, Created)
DT  01-AUG-1992 (Rel. 23, Last sequence update)
DT  15-SEP-2003 (Rel. 42, Last annotation update)
DE  Glucosyltransferase-I precursor (EC 2.4.1.5) (GTF-I) (Dextranucrase)
DE  (Sucrose 6-glucosyltransferase).
OS  Streptococcus downei (Streptococcus sobrinus).
OC  Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC  Streptococcus.
OX  NCBI_TaxID=1317;
RN  [1]
RP  SEQUENCE FROM N.A.
RC  STRAIN=6715 / Serotype G;
RX  MEDLINE=91123227; Pubmed=1704006;
RA  Abo H., Matsumura T., Kodama T., Ohta H., Fukui K., Kato K.,
RA  Kagawa H.;
RT  "Peptide sequences for sucrose splitting and glucan binding within
RT  Streptococcus sobrinus glucosyltransferase (water-insoluble glucan
RT  synthetase).";
RL  J. Bacteriol. 173:989-996(1991).
CC  -1- FUNCTION: PRODUCTION OF EXTRACELLULAR GLUCANS, THAT ARE THOUGHT
CC  TO PLAY A KEY ROLE IN THE DEVELOPMENT OF THE DENTAL PLAQUE BECAUSE
CC  OF THEIR ABILITY TO ADHERE TO SMOOTH SURFACES AND MEDIATE THE
CC  AGGREGATION OF BACTERIAL CELLS AND FOOD DEBRIS.
CC  -1- CATALYTIC ACTIVITY: Sucrose + ((1,6)-alpha-D-glucosyl) (N-1) = D-
CC  fructose + ((1,6)-alpha-D-glucosyl) (N-1).
CC  -1- SUBCELLULAR LOCATION: Secreted.
CC  -1- DISEASE: DENTAL CARIES.
CC  -1- MISCELLANEOUS: GTF-I SYNTHESIZES WATER-INSOLUBLE GLUCANS (ALPHA
CC  1,3-LINKED GLUCOSE AND SOME 1,6 LINKAGES), GTF-S SYNTHESIZES
CC  WATER-SOLUBLE GLUCANS (ALPHA 1,6-GLUCOSE). GTF-SI SYNTHESIZES BOTH
CC  FORMS OF GLUCANS.
CC  -1- SIMILARITY: TO OTHER GLUCOSYLTRANSFERASES AND SOME TO A GLUCAN-
CC  BINDING PROTEIN FROM S. MUTANS.
CC  -1- SIMILARITY: Contains 16 cell wall binding repeats.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC  or send an email to license@sib-sib.ch).
CC  -----
CC  EMBL; D90213; BA014241.1; -
CC  InterPro; IPR002479; CW binding.
CC  InterPro; IPR003118; Glyco_hydro_70.
CC  Pfam; PF01473; CW_binding_I_13.
CC  Pfam; PF03324; Glyco_hydro_70; 1.
CC  Trasnferase; Glucosyltransferase; Signal; Repeat; Dental caries.
CC  SIGNAL 1 38 POTENTIAL.
CC  CHAIN 39 1592 GLUCOSYLTRANSFERASE-I.
CC  FT 39 1592 CATALYTIC (APPROXIMATE).
CC  FT DOMAIN 39 1044

```



```

CC -1- MISCELLANEOUS: GTF-1 SYNTHESIZES WATER-INSOLUBLE GLUCANS (ALPHA
CC 1,3-LINKED GLUCOSE AND SOME 1,6 LINKAGES). GTF-S SYNTHESIZES
CC WATER-SOLUBLE GLUCANS (ALPHA 1,6-GLUCOSE). GTF-SI SYNTHESIZES BOTH
CC FORMS OF GLUCANS.
CC -1- SIMILARITY: TO OTHER GLUCOSYLTRANSFERASES AND SOME TO A GLUCAN-
CC BINDING PROTEIN FROM S. MUTANS.
CC -1- SIMILARITY: Contains 10 cell wall binding repeats.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M17361; AAA8588.1; -
DR EMBL; D88651; BAA26101.1; -
DR EMBL; D88654; BAA26105.1; -
DR EMBL; D88657; BAA26109.1; -
DR EMBL; D88660; BAA26113.1; -
DR EMBL; D89977; BAA26119.1; -
DR EMBL; AE014940; AAN58705.1; -
DR InterPro; IPR002479; CM binding.
DR InterPro; IPR003318; Glyco hydro_70.
DR Pfam; PF01473; CM-binding_1.1.
DR Pfam; PF02324; Glyco_hydro_70; 1.
KM Transferrase; Glycosyltransferase; signal; Repeat; Dental caries;
KM Complete proteome.
FT SIGNAL 1 34
FT CHAIN 35 1476
FT DOMAIN 35 1051
FT DOMAIN 1097 1476
FT REPEAT 1097 1130
FT REPEAT 1161 1470
FT REPEAT 1225 1275
FT REPEAT 1290 1340
FT REPEAT 1355 1405
FT REPEAT 1420 1470
FT REPEAT 62 62
FT REPEAT 65 65
FT REPEAT 68 68
FT VARIANT 78 78
FT VARIANT 86 86
FT VARIANT 89 89
FT VARIANT 168 168
FT VARIANT 276 276
FT VARIANT 399 399
FT VARIANT 474 474
FT VARIANT 512 512
FT VARIANT 519 519
FT VARIANT 701 701
FT VARIANT 708 708
FT VARIANT 938 938
FT VARIANT 952 952
FT VARIANT 963 964
FT VARIANT 968 970
FT VARIANT 1086 1086
FT VARIANT 1158 1158
FT VARIANT 1163 1163
FT VARIANT 1168 1168
FT VARIANT 1182 1182
FT VARIANT 1234 1234
FT VARIANT 1263 1263
FT VARIANT 1264 1264

```

```

FT FT AND MT8148).
FT FT S -> G (IN STRAINS GS-5, MT4239, MT4467
FT FT AND MT8148).
FT FT H -> Y (IN STRAINS GS-5 AND MT4467).
FT FT Y -> Y (IN STRAINS GS-5, MT4239, MT4467
FT FT AND MT8148).
FT FT S -> G (IN STRAINS GS-5, MT4239, MT4467
FT FT AND MT8148).
FT FT Y -> H (IN STRAIN MT4467).
FT FT R -> A (IN REF. 1).
FT FT ADDYRVASTAPSTDGK -> LKMFALRLARPHQOMA
FT FT (IN REF. 1).
FT FT H -> L (IN REF. 1).
SQ SEQUENCE 1476 AA; 165846 MW; 906E09F731B4CBCE CRC64;

Query Match
Best Local Similarity 90.8%; Score 108; DB 1; Length 1476;
Matches 18; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 ANNVHSIVEAGSDNDPTVYIHD 21
DB 481 ANDHLSILEAWSNDPTVYIHD 501

RESULT 4
GTFC STRMU STANDARD; PRT; 1455 AA.
ID PF1470; 069382; 069385; 069388; 069391; 069397; P05427;
AC 01-NOV-1988 (Rel. 09, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Glucosyltransferase-SI precursor (EC 2.4.1.5) (GTF-SI)
DE (Dextranucrase) (Sucrose 6-glucosyltransferase).
GN GTFC OR SMU.1005.
OS Streptococcus mutans.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1309;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=GS-5;
RX MEDLINE=89137980; PubMed=2976010;
RA Ueda S., Shiroza T., Kuramitsu H.K.;
RT "Sequence analysis of the gtfC gene from Streptococcus mutans GS-5.";
RL Gene 69:101-109(1988).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=MT4239 / Serotype C, MT4245 / Serotype E, MT4251 / Serotype F,
RX MT4467 / Serotype B, and MT8148 / Serotype C;
RX MEDLINE=98231643; PubMed=9570124;
RA Fujiwara T., Terao Y., Hoshino T., Kawabata S., Ooshima T., Sobue S.,
RA Kimura S., Hamada S.;
RT "Molecular analyses of glucosyltransferase genes among strains of
RT Streptococcus mutans.";
RL FEMS Microbiol. Lett. 161:331-336(1998).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=UA159 / ATCC 700610 / Serotype C;
RX MEDLINE=22295063; PubMed=12397185;
RA Ajdic D., Mcnehan W.M., McLaughlin R.E., Savic G., Chang J.,
RA Carson W.B., Primeaux C., Tian R., Kenton S., Jia H., Lin S., Qian Y.,
RA Li S., Zhu H., Najjar F., Lai H., White U., Roe B.A., Ferretti J.J.;
RT "Genome sequence of Streptococcus mutans UA159, a cariogenic dental
RT pathogen.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:14434-14439(2002).
RN [4]
RP SEQUENCE OF 1-349 FROM N.A.
RC STRAIN=GS-5;
RX MEDLINE=87308013; PubMed=3040685;
RA Shiroza T., Ueda S., Kuramitsu H.K.;
RT "Sequence analysis of the gtfB gene from Streptococcus mutans.";
RL J. Bacteriol. 169:4263-4270(1987).
CC -1- FUNCTION: PRODUCTION OF EXTRACELLULAR GLUCANS, THAT ARE THOUGHT

```

TO PLAY A KEY ROLE IN THE DEVELOPMENT OF THE DENTAL PLAQUE BECAUSE OF THEIR ABILITY TO ADHERE TO SMOOTH SURFACES AND MEDIATE THE AGGREGATION OF BACTERIAL CELLS AND FOOD DEBRIS.

- CATALYTIC ACTIVITY: Sucrose + ((1,6)-alpha-D-glucosyl) (N) = D-fructose + ((1,6)-alpha-D-glucosyl) (N+1).

- SUBCELLULAR LOCATION: Secreted.

- DISEASE: DENTAL CARIES.

- MISCELLANEOUS: GTF-I SYNTHESIZES WATER-INSOLUBLE GLUCANS (ALPHA 1,3-LINKED GLUCOSE AND SOME 1,6 LINKAGES), GTF-S SYNTHESIZES WATER-SOLUBLE GLUCANS (ALPHA 1,6-GLUCOSE). GTF-SI SYNTHESIZES BOTH FORMS OF GLUCANS.

- SIMILARITY: TO OTHER GLUCOSYLTRANSFERASES AND SOME TO A GLUCAN-BINDING PROTEIN FROM S. MUTANS.

- SIMILARITY: Contains 5 cell wall binding repeats.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; M2054; AAA88592.1; -
EMBL; D8652; BAA26102.1; -
EMBL; D8655; BAA26106.1; -
EMBL; D8658; BAA26110.1; -
EMBL; D8661; BAA26114.1; -
EMBL; D8978; BAA26120.1; -
EMBL; A014940; AAN58706.1; -
EMBL; M17361; AAA8589.1; -
InterPro: IPR002479; CW binding.
DR InterPro: IPR003318; Glyco_hydro_70.
DR Pfam: PF01473; CW binding_1; 1.
DR Pfam: PF02324; Glyco_hydro_70; 1.
Transferrase; Glycosyltransferase; Signal; Repeat; Dental caries;
KW Complete proteome.

FT SIGNAL 34
FT CHAIN 35 1455
FT DOMAIN 35 1050
FT DOMAIN 1126 1455
FT DOMAIN 1126 1455
FT REPEAT 1126 1159
FT REPEAT 1169 1200
FT REPEAT 1227 1238
FT REPEAT 1253 1303
FT REPEAT 1318 1330
FT VARIANT 21 21
FT VARIANT 81 81
FT VARIANT 106 106
FT VARIANT 116 116
FT VARIANT 126 126
FT VARIANT 150 151

FT VARIANT 256 256
FT VARIANT 425 425
FT VARIANT 519 519
FT VARIANT 538 538
FT VARIANT 545 545
FT VARIANT 597 597

FT VARIANT 600 600
FT VARIANT 601 601
FT VARIANT 614 614
FT VARIANT 727 727
FT VARIANT 734 734
FT VARIANT 964 964
FT VARIANT 1113 1113
FT VARIANT 1118 1118
FT VARIANT 1204 1204
FT VARIANT 1208 1208

GLUCOSYLTRANSFERASE-ST.
CATALYTIC (APPROXIMATE).
GLUCAN-BINDING (APPROXIMATE).
2.4.A.1 C AND 1 AC REPEATS.
A REPEAT.
C REPEAT.
AC REPEAT.
A REPEAT (INCOMPLETE).
V -> I (IN STRAIN GS-5).
P -> L (IN STRAIN MT4239).
D -> V (IN STRAIN GS-5).
S -> A (IN STRAIN GS-5 AND MT4467).
A -> T (IN STRAIN GS-5).
SR -> PK (IN STRAINS GS-5, MT4239 AND MT4467).
A -> V (IN STRAINS GS-5 AND MT4467).
R -> N (IN STRAIN MT4251).
Y -> D (IN STRAINS MT4245 AND MT4251).
R -> K (IN STRAINS MT4245 AND MT4251).
Y -> F (IN STRAINS MT4245 AND MT4251).
N -> D (IN STRAINS MT4245, MT4251, MT4467 AND MT8148).
R -> K (IN STRAINS MT4245, MT4251, MT4467 AND MT8148).
A -> T (IN STRAIN GS-5).
M -> I (IN STRAIN MT8148).
T -> I (IN STRAIN MT8148).
L -> V (IN STRAIN MT8148).
L -> F (IN STRAIN MT4239).
A -> Y (IN STRAIN MT4239).
A -> T (IN STRAIN MT4239).
I -> V (IN STRAINS GS-5, MT4239, MT4467 AND MT8148).
V -> I (IN STRAIN MT8148).

FT VARIANT 1292 1294 DGH -> NGY (IN STRAINS GS-5, MT4467 AND MT8148).
FT VARIANT 1305 1369 MISSING (IN STRAIN MT4245).
FT VARIANT 1326 1326 I -> V (IN STRAIN MT8148).
FT VARIANT 1331 1331 T -> A (IN STRAINS GS-5, MT4239, MT4467 AND MT8148).
FT VARIANT 1377 1377 R -> K (IN STRAIN MT8148).
FT VARIANT 1398 1398 V -> I (IN STRAIN MT8148).
FT VARIANT 1424 1424 D -> N (IN STRAIN MT4239).
FT VARIANT 1439 1439 V -> I (IN STRAINS MT4239 AND MT8148).
FT VARIANT 1444 1444 S -> P (IN STRAIN MT8148).
FT CONFLICT 1337 1455 ORLYKSNQVQAKGLITERKRIKYDPNSGNEVNRNRYR TSGNMWYFPGNDGYALIGHVGEGRVYFEDNGVRYASHD QRNMWDYDRDRRCGSSSAVRFPSRNGFDFNFRF -> HASTLSLWFRRLRRESSLSQSVYVGNMTMLIEMKFLVIM (IN REF. 1).

SEQUENCE 1455 AA; 162965 MW; 3CB455A924FEC86 CRC64;
Query Match 83.2%; Score 99; DB 1; Length 1455;
Best Local Similarity 81.0%; Pred. No. 1.6e-07;
Matches 17; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

1 ANNHVIVEAMSDNDTPYLHD 21
507 ANDHSLIEAMSYNDTPYLHD 527

RESULT 5
GTFD STRMU STANDARD; PRT; 1462 AA.
ID GTFD STRMU 069383; 069386; 069389; 069392; 069398;
AC P49331, 069383; 069386; 069389; 069392; 069398;
DT 01-FEB-1996 (Rel. 33, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Glucosyltransferase-8 precursor (BC 2.4.1.5) (GTF-S) (Dextranucrase)
DE (Sucrose 6-glucosyltransferase).
GN GTFD OR SMU.910.
OS Streptococcus mutans.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxId=1309;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=GS-5
RX MEDLINE=91100958; PubMed=2148600;
RA Honda O., Kato C., Kuramitsu H.K.;
RT "Nucleotide sequence of the Streptococcus mutans gtfD gene encoding the glucosyltransferase-8 enzyme."
RL J. Gen. Microbiol. 136:2099-2105 (1990).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=MT4239 / Serotype C, MT4245 / Serotype E, MT4251 / Serotype F,
RX MT4467 / Serotype E, and MT8148 / Serotype C;
RA MEDLINE=9821643; PubMed=9570124;
RX Fujiwara T., Terao Y., Hoshino T., Kawabata S., Ooshima T., Sobue S., Kimura S., Hamada S.;
RT "Molecular analyses of glucosyltransferase genes among strains of Streptococcus mutans."
RL FEWS Microbiol. Lett. 161:331-336 (1998).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=UAI59 / ATCC 700610 / Serotype C;
RX MEDLINE=22295063; PubMed=12397186;
RA Ajdic D., Moshan W.M., McLaughlin R.E., Savic G., Chang J., Carson M.B., Primeaux C., Tjian R., Kenton S., Jia H., Lin S., Qian Y., Li S., Zhu H., Najjar F., Lai H., White J., Roe B.A., Ferretti J.J.;
RT "Genome sequence of Streptococcus mutans UAI59, a cariogenic dental pathogen."
RL Proc. Natl. Acad. Sci. U.S.A. 99:14434-14439 (2002).
CC -1- FUNCTION: PRODUCTION OF EXTRACELLULAR GLUCANS, THAT ARE THOUGHT TO PLAY A KEY ROLE IN THE DEVELOPMENT OF THE DENTAL PLAQUE BECAUSE OF THEIR ABILITY TO ADHERE TO SMOOTH SURFACES AND MEDIATE THE AGGREGATION OF BACTERIAL CELLS AND FOOD DEBRIS.

CC -1- CATALYTIC ACTIVITY: Sucrose + {(1,6)-alpha-D-glucosyl}(N) = D-fructose + {(1,6)-alpha-D-glucosyl}(N+1).

CC -1- SUBCELLULAR LOCATION: Secreted.

CC -1- DISEASE: DENTAL CARIES.

CC -1- MISCELLANEOUS: GTF-1 SYNTHESIZES WATER-INSOLUBLE GLUCANS (ALPHA 1,3-LINKED GLUCOSE AND SOME 1,6 LINKAGES), GTF-S SYNTHESIZES WATER-SOLUBLE GLUCANS (ALPHA 1,6-GLUCOSE). GTF-SI SYNTHESIZES BOTH FORMS OF GLUCANS.

CC -1- SIMILARITY: TO OTHER GLUCOSYLTRANSFERASES AND SOME TO A GLUCAN-BINDING PROTEIN FROM S. MUTANS.

CC -1- SIMILARITY: Contains 6 cell wall binding repeats.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC EMBL: M29286; AAA26895.1; -

DR EMBL: D88653; BAA26103.1; -

DR EMBL: D88656; BAA26107.1; -

DR EMBL: D88659; BAA26111.1; -

DR EMBL: D88662; BAA26115.1; -

DR EMBL: D89979; BAA26121.1; -

DR EMBL: AF014932; AAN58619.1; -

DR InterPro: IPR002479; CW_binding_70.

DR InterPro: IPR003318; Glyco_hydro_70; 1.

DR Pfam: PF02324; Glyco_hydro_70; 1.

KW Transferase; Glycosyltransferase; Signal; Repeat; Dental caries; Complete proteome.

FT SIGNAL 1 ?

FT CHAIN ? 1462

FT DOMAIN 1232 1423

FT REPEAT 1232 1295

FT REPEAT 1296 1359

FT REPEAT 1360 1423

FT REPEAT 10 10

FT VARIANT 19

FT VARIANT 58

FT VARIANT 68

FT VARIANT 81

FT VARIANT 113

FT VARIANT 122

FT VARIANT 132

FT VARIANT 135

FT VARIANT 137

FT VARIANT 202

FT VARIANT 255

FT VARIANT 275

FT VARIANT 288

FT VARIANT 301

FT VARIANT 313

FT VARIANT 317

FT VARIANT 328

FT VARIANT 350

FT VARIANT 628

FT VARIANT 688

FT VARIANT 726

FT VARIANT 730

FT VARIANT 762

FT VARIANT 964

FT VARIANT 1019

CC POTENTIAL.

CC GLUCOSYLTRANSFERASE-S.

CC 3 X 63 AA APPROXIMATE TANDEM REPEATS.

CC 1.

CC 2.

CC 3.

CC Y->H (IN STRAINS GS-5, MT4239, MT4245, MT4251, MT4467 AND MT8148).

CC I->V (IN STRAINS GS-5, MT4239, MT4245, MT4251, MT4467 AND MT8148).

CC K->E (IN STRAIN MT4467).

CC A->S (IN STRAINS MT4239 AND MT4245).

CC T->T (IN STRAINS MT4251 AND MT8148).

CC T->I (IN STRAINS MT4239, MT4245 AND MT8148).

CC A->V (IN STRAINS GS-5 AND MT4467).

CC A->S (IN STRAIN MT4245).

CC A->T (IN STRAINS GS-5, MT4239, MT4245, MT4251, MT4467 AND MT8148).

CC D->N (IN STRAIN MT4239).

CC D->L (IN STRAIN MT8148).

CC E->D (IN STRAINS MT4239, MT4245 AND MT4251).

CC D->N (IN STRAINS MT4239, MT4245 AND MT4251).

CC Q->H (IN STRAIN MT4245).

CC D->N (IN STRAINS MT4239 AND MT4251).

CC E->K (IN STRAIN MT4239).

CC V->P (IN STRAIN MT4239).

CC F->L (IN STRAINS MT4239, MT4251 AND MT4467).

CC KKKYQ->EKEYTL (IN STRAIN MT4251).

CC A->S (IN STRAIN MT4239).

CC TDGSEA->ADKGN (IN STRAIN MT4251).

CC TDGSEA->ADKGN (IN STRAINS MT4239 AND MT4245).

CC T->A (IN STRAINS GS-5, MT4239, MT4245, MT4251, MT4467 AND MT8148).

CC D->Y (IN STRAIN MT4251).

CC E->K (IN STRAINS MT4245 AND MT4251).

FT VARIANT 1059 1060 IG->IR (IN STRAIN MT4251).

FT VARIANT 1060 1060 G->R (IN STRAIN MT4245).

FT VARIANT 1080 1080 G->R (IN STRAIN MT4239).

FT VARIANT 1142 1142 H->Q (IN STRAIN GS-5).

FT VARIANT 1198 1198 S->N (IN STRAIN MT4239).

FT VARIANT 1220 1220 Y->C (IN STRAINS MT4251 AND MT4467).

FT VARIANT 1280 1280 F->L (IN STRAIN MT4467).

FT VARIANT 1282 1282 Q->P (IN STRAIN MT4245).

FT VARIANT 1290 1290 K->T (IN STRAIN MT4245).

FT VARIANT 1311 1311 N->D (IN STRAIN MT4245).

FT VARIANT 1403 1403 G->D (IN STRAINS GS-5 AND MT4467).

FT VARIANT 1425 1425 G->R (IN STRAIN GS-5).

FT VARIANT 1449 1449 R->K (IN STRAIN MT4467).

FT VARIANT 1462 1462 RYDNGSNVYKVTLANGRIGIDPWGIARYV -> VY R (IN REF. 1).

FT CONFLICT 1428 1462

FT SEQUENCE 1462 AA; 163387 MW; CE4A279C4D708645 CRC64;

QY Query Match 62.2%; Score 74; DB 1; Length 1462; Best Local Similarity 66.7%; Pred. No. 0.0015; Matches 14; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Db 495 ANHSTLEAMSDNDPQYKMD 515

QY 1 ANHSTLEAMSDNDPQYKMD 21

DB 495 ANHSTLEAMSDNDPQYKMD 515

RESULT 6

GFES STRDO STANDARD; PRT; 1365 AA.

AC P29336;

DT 01-DEC-1992 (Rel. 24, Created)

DT 01-DEC-1992 (Rel. 24, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Glucosyltransferase-S precursor (EC 2.4.1.5) (GTF-S) (Dextranucrase) (Sucrose 6-glucosyltransferase).

DE GTFES.

OS Streptococcus downei (Streptococcus sobrinus).

OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae; Streptococcus.

OX NCBI_TaxID=1317;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=MFE28;

RC MEDLINE=90316665; PubMed=2142479;

RA "Almore K.S., Russell R.R., Ferretti J.J.;

RT "Analysis of the Streptococcus downei gtfS gene, which specifies a glucosyltransferase that synthesizes soluble glucans.";

RT Infect. Immun. 58:2452-2458 (1990).

RL -1- FUNCTION: PRODUCTION OF EXTRACELLULAR GLUCANS, THAT ARE THOUGHT TO PLAY A KEY ROLE IN THE DEVELOPMENT OF THE DENTAL PLAQUE BECAUSE OF THEIR ABILITY TO ADHERE TO SMOOTH SURFACES AND MEDIATE THE AGGREGATION OF BACTERIAL CELLS AND FOOD DEBRIS.

CC -1- CATALYTIC ACTIVITY: Sucrose + {(1,6)-alpha-D-glucosyl}(N) = D-fructose + {(1,6)-alpha-D-glucosyl}(N+1).

CC -1- ENZYME REGULATION: GLUCAN SYNTHESIS BY GTF-S IS INDEPENDENT OF PRIMER GLUCAN UNLIKE GTF-1.

CC -1- DISEASE: DENTAL CARIES.

CC -1- MISCELLANEOUS: GTF-S SYNTHESIZES WATER-SOLUBLE GLUCANS (ALPHA 1,6-GLUCOSE).

CC -1- SIMILARITY: TO OTHER GLUCOSYLTRANSFERASES AND SOME TO A GLUCAN-BINDING PROTEIN FROM S. MUTANS.

CC -1- SIMILARITY: Contains 10 cell wall binding repeats.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC EMBL: M30943; AAA26898.1; -

DR InterPro: IPR002479; CW_binding.

```

DR   InterPro: IPR003318; Glyco_hydro_70.
DR   Pfam: PF01473; CW binding 1; 8.
DR   Pfam: PF02324; Glyco_hydro_70; 1.
DR   Pfam: PF02324; Glycosyltransferase; Signal; Repeat; Dental caries.
KW   Transferase; Glycosyltransferase; Signal; Repeat; Dental caries.
FT   SIGNAL 1 36
FT   CHAIN 1 36
FT   REPEAT 157 177
FT   REPEAT 178 197
FT   DOMAIN 198 1061
FT   REPEAT 1062 1082
FT   REPEAT 1083 1102
FT   REPEAT 1150 1169
FT   REPEAT 1170 1190
FT   REPEAT 1225 1243
FT   REPEAT 1289 1308
FT   REPEAT 1309 1328
FT   REPEAT 1331 1352
SQ   SEQUENCE 1365 AA; 151590 MW; 167296B5A2E6C476 CRC64;

Query Match 56.3%; Score 67; DB 1; Length 1365;
Best Local Similarity 57.1%; Pred. No. 0.018;
Matches 12; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

OY 1 ANNHVSIWEAASNDNDPTFLHD 21
Db 467 AIDHLSIEAASGNDNDYVD 487

RESULT 7
ID   UL16_EBV STANDARD; PRT; 336 AA.
AC   P03221;
DT   21-JUL-1986 (Rel. 01, Created)
DT   21-JUL-1986 (Rel. 01, Last sequence update)
DT   15-DEC-1998 (Rel. 37, Last annotation update)
DE   BGLF2 protein.
GN   BGLF2.
OS   Epstein-Barr virus (strain B95-8) (Human herpesvirus 4).
OC   Viruses; daDNA viruses; no RNA stage; Herpesviridae;
OC   Gammaherpesvirinae; Lymphocryptovirus.
OX   NCBI_TaxID=10377;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=84270667; PubMed=6087149;
RX   Baer R., Bankier A.T., Biggin M.D., Deininger P.L., Farrell P.J.,
RX   Gibson T.J., Hatfull G., Hudson G.S., Satchwell S.C., Seguin C.,
RX   Tuffnell P.S., Barrett B.G.;
RA   "DNA sequence and expression of the B95-8 Epstein-Barr virus genome.";
RA   Nature 310:207-211 (1984).
RL   [2]
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=92113548; PubMed=1662696;
RX   Chen M.R., Hsu T.Y., Lin S.W., Chen J.Y., Yang C.S.;
RA   "Cloning and characterization of cDNA clones corresponding to
RA   transcripts from the BamHI G region of the Epstein-Barr virus genome
RA   and expression of BGLF2.";
RT   J. Gen. Virol. 72:3047-3055 (1991).
RL   J. Gen. Virol. 72:3047-3055 (1991).
CC   -1- SIMILARITY: BELONGS TO FAMILY THAT GROUPS TOGETHER HSV-1 UL16,
CC   HSV-6 ORF1R, EBV-1 46, HCMV UL94, EBV BDLF2, AND VZV 44.
CC   -----
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   -----
CC   EMBL: M60514; AAA45871.1; -
CC   EMBL: V01555; CAA24831.1; -
CC   EMBL: S77132; AAB21113.1; -
CC   PIR: C43044; Q0BEA0.
CC   InterPro: IPR004286; UL16_UL94.

```

```

DR   Pfam: PF03044; UL16_UL94; 1.
KW   Late protein.
SQ   SEQUENCE 336 AA; 36888 MW; 840937A416D584C CRC64;

Query Match 41.2%; Score 49; DB 1; Length 336;
Best Local Similarity 41.2%; Pred. No. 2.5;
Matches 7; Conservative 6; Mismatches 4; Indels 0; Gaps 0;

OY 1 ANNHVSIWEAASNDNDPT 17
Db 216 AGAHVNIKRGHTEDDSP 232

RESULT 8
ID   G3PC_LEIME STANDARD; PRT; 330 AA.
AC   Q01558;
DT   01-APR-1993 (Rel. 25, Created)
DT   01-FEB-1994 (Rel. 28, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Glyceralddehyde 3-phosphate dehydrogenase, cytosolic (EC 1.2.1.12)
DE   (GAPDH).
GN   GAPC.
OS   Leishmania mexicana.
OC   Eukaryota; Euklenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.
OX   NCBI_TaxID=5665;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   STRAIN=Sep. Mexicana;
RX   MEDLINE=93063042; PubMed=1435864;
RX   Hannaert V., Blaauw M., Kohl L., Allert S., Opperdoes F.R.,
RX   Michels P.A.M.;
RA   "Molecular analysis of the cytosolic and glycosomal glyceralddehyde-3-
RA   phosphate dehydrogenase in Leishmania mexicana.";
RT   Mol. Biochem. Parasitol. 55:115-126 (1992).
RL   -1- CATALYTIC ACTIVITY: D-glyceraldehyde 3-phosphate + phosphate +
CC   NAD(+) = 3-phospho-D-glyceroyl phosphate + NADH.
CC   -1- PATHWAY: Second phase of glycolysis; first step.
CC   -1- SUBUNIT: Homotetramer.
CC   -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC   -1- SIMILARITY: Belongs to the glyceralddehyde 3-phosphate
CC   dehydrogenase family.
CC   -----
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   -----
CC   EMBL: X65220; CAA46323.1; -
CC   PIR: B48445; B48445.
CC   HSBP: P06977; IGAD.
CC   InterPro: IPR000173; GAP dehydrogenase.
DR   InterPro: IPR006424; GAPDH-I.
DR   Pfam: PF00044; spdh; 1.
DR   Pfam: PF02800; spdh C; 1.
DR   PRINTS: PR00078; G3PDHGRNASE.
DR   TIGRfam: TIGR01534; GAPDH-I; 1.
DR   PROSITE: PS00071; GAPDH; 1.
KW   Glycolysis; Oxidoreductase; NAD.
FT   INIT MET 0
FT   BINDING 148 148
FT   ACT SITE 175 175
SQ   SEQUENCE 330 AA; 35511 MW; EDAB6DBE8A207F1E CRC64;

Query Match 40.3%; Score 48; DB 1; Length 330;
Best Local Similarity 42.1%; Pred. No. 3.5;
Matches 8; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

OY 2 NNHVSIWEAASNDNDPTFLH 20

```

DB 300 NDHFVKLVSWYDNETGYSH 318

RESULT 9

YMS8_YEAST STANDARD; PRT; 313 AA.

AC 003695;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Hypothetical 35.0 kDa protein in PER2-HFAl intergenic region.
GN YMR206W OR YMR325.07.
OS *Saccharomyces cerevisiae* (Baker's yeast).
OC *Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes; Saccharomycetales; Saccharomycetaceae; Saccharomyces.*
OX NCBI_TaxID=4932;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=5288C / AB972;
RX PubMed=9169872;
RA Bowman S., Churcher C.M., Badcock K., Brown D., Chillingworth T., Connor R., Dedman K., Devlin K., Gentles S., Hamlin N., Hunt S., Jagsels K., Lyle G., Moule S., Odell C., Pearson D., Rajadream M.A., Rice P., Skelton J., Walsh S., Whitehead S., Barrett B.G.;
RA "The nucleotide sequence of *Saccharomyces cerevisiae* chromosome XII";
RT Nature 387:90-93 (1997).
RL -1- SIMILARITY: SOME, TO YEAST YMR014W.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - CC the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC EMBL; 248755; CAA8648.1; -;
DR PIR; S59448; S59448.
DR SGD; S0004819; YMR206W.
KW Hypothetical protein.
FT DOMAIN 3
FT DOMAIN 146 149 POLY-SER.
FT DOMAIN 246 252 POLY-GLN.
FT DOMAIN 313 AA; 35017 MW; 9D92BDE98257770 CRC64;
SQ SEQUENCE

Query Match 38.7%; Score 46; DB 1; Length 313;
Best Local Similarity 47.4%; Pred. No. 6.9;
Matches 9; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

OY 2 NNHVSIVAMSDNDTPYLH 20
DB 125 NNHQTARFSDSHPSLH 143

RESULT 10
G3P MYCGE STANDARD; PRT; 337 AA.

AC P47543;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Glyceraldehyde 3-phosphate dehydrogenase (EC 1.2.1.12) (GAPDH).
GN GAP3 OR GAP3 OR MG301.
OS *Mycoplasma genitalium*.
OC *Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.*
OX NCBI_TaxID=2097;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 33530 / G-37;
RX MEDLINE=96026346; PubMed=7569993;
RA Frazer C.M., Gocayne J.D., White O., Adams M.D., Clayton R.A., Fleischmann R.D., Bult C.J., Kerlavage A.R., Sutton G., Kelley J.M.,

RA Fritchman J.L., Weidman J.F., Small K.V., Sandusky M., Fuhrmann J.L., Nguyen D.T., Uterback T.R., Saudek D.M., Phillips C.A., Merrick J.M., Tomb J.F., Dougherty B.A., Bott K.F., Hu P.-C., Lucier T.S., Peterson S.N., Smith H.O., Hutchison C.A. III, Venter J.C.;
RA "The minimal gene complement of *Mycoplasma genitalium*."; Science 270:397-403 (1995).

RT [2]
RT SEQUENCE OF 1-81 AND 279-337 FROM N.A.
RP STRAIN=ATCC 33530 / G-37;
RC MEDLINE=94075230; PubMed=8253680;
RX Peterson S.N., Hu P.-C., Bott K.F., Hutchison C.A. III, Venter J.C.;
RT "A survey of the *Mycoplasma genitalium* genome by using random sequencing";
RT J. Bacteriol. 175:7918-7930 (1993).
RL -1- CATALYTIC ACTIVITY: D-glyceraldehyde 3-phosphate + phosphate + NAD(+) = 3-phospho-D-glyceroyl phosphate + NADH.
CC -1- PATHWAY: Second phase of glycolysis; first step.
CC -1- SUBUNIT: Homotetramer (By similarity).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: Belongs to the glyceraldehyde 3-phosphate dehydrogenase family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - CC the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC EMBL; U9710; AAC71523.1; -;
DR EMBL; U02213; AAD12507.1; -;
DR EMBL; U02178; AAD12463.1; -;
DR PIR; C64233; C64233.
DR HSSP; P17721; 1HDG.
DR TIGR; MG301; -;
DR Interpro: IPR000173; GAP_dhhydrogenase.
DR Interpro: IPR006424; GAPDH-I.
DR Pfam; PF00044; gpdh_1.
DR Pfam; PF02800; gpdh_C_1.
DR PRINTS; PR00078; G3PDHDSGNASE.
DR TIGRFAMs; TIGR01534; GAPDH-I; 1.
DR PROSITE; PS00071; GAPDH; 1.
KW Glycolysis; Oxidoreductase; NAD; Complete proteome.
FT BINDING 157 157 GLYCERALDEHYDE 3-PHOSPHATE (BY SIMILARITY).
FT ACT_SITE 184 184 ACTIVATES THIOL GROUP DURING CATALYSIS (BY SIMILARITY).
FT FT
SQ SEQUENCE 337 AA; 37097 MW; FA1EA19668706B CRC64;

Query Match 37.8%; Score 45; DB 1; Length 337;
Best Local Similarity 53.8%; Pred. No. 11;
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

OY 8 YEAMSDNDTPYLH 20
DB 313 VYAMYNDSYVH 325

RESULT 11
SILF MOUSE STANDARD; PRT; 569 AA.

AC 092063;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Sialic acid binding Ig-like lectin-F precursor (msiglec-F).
GN SIGLECF.
OS *Mus musculus* (Mouse).
OC *Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.*
OX NCBI_TaxID=10090;
RN (1)

```

RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6;
RX MEDLINE=21576254; PubMed=11579105;
RA Angata T., Hingorani R., Varki N.M., Varki A.;
RT "Cloning and characterization of a novel mouse Siglec, mSiglec-F;
RT differential evolution of the mouse and human (CD33) Siglec-3-related
RT gene clusters."
RL J. Biol. Chem. 276:45128-45136(2001).
CC -1- FUNCTION: Putative adhesion molecule that mediates sialic-acid
CC -1- dependent binding to cells. Preferentially binds to alpha2,3-
CC linked sialic acid. The sialic acid recognition site may be masked
CC by cis interactions with sialic acids on the same cell surface.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- TISSUE SPECIFICITY: Predominantly expressed by immature
CC monocytic/myeloid lineage cells in bone marrow. Also found at
CC lower levels in mature neutrophils and monocytes.
CC -1- DOMAIN: Contains an intracytoplasmic motif referred as
CC immunoreceptor tyrosine-based inhibitor motif (ITIM). This motif
CC is involved in downmodulation of cellular functions as the
CC termination of the immune response.
CC -1- SIMILARITY: Belongs to the immunoglobulin superfamily. SIGLEC
CC (SIALIC ACID BINDING Ig-LIKE LECTIN) SUBFAMILY.
CC -1- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.
CC -1- SIMILARITY: Contains 2 immunoglobulin-like C2-type domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AF293371; AAL1043.1; -.
DR InterPro: IPR007110; Ig-like.
DR InterPro: IPR003598; Ig_C2.
DR InterPro: IPR003006; Ig_MHC.
DR Pfam: PF00047; Ig_2.
DR SMART: SM00408; IGC2; 1.
DR PROSITE: PS50835; IG_LIKE; 2.
DR Cell adhesion; Lectin; Transmembrane; Signal; Glycoprotein;
KW Immunoglobulin domain; Repeat.
FT SIGNAL 1 16
FT CHAIN 17 569
FT DOMAIN 17 439
FT TRANSMEM 440 460
FT DOMAIN 461 569
FT DOMAIN 18 116
FT DOMAIN 139 224
FT DOMAIN 229 324
FT SITE 536 541
FT SITE 559 564
FT SITE 35 163
FT DISULFID 35 163
FT DISULFID 40 96
FT DISULFID 157 206
FT DISULFID 265 308
FT CARBOHYD 95 95
FT CARBOHYD 151 151
FT CARBOHYD 200 200
FT CARBOHYD 203 203
FT CARBOHYD 369 369
FT CARBOHYD 372 372
FT CARBOHYD 387 387
SQ SEQUENCE 569 AA; 61476 MW; 8093838090484FC1 CRC64;

Query Match 37.0%; Score 44; DB 1; Length 569;
Best Local Similarity 57.1%; Pred. No. 30;
Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

```

```

RESULT 12
Y33C YEAST STANDARD; PRT; 1122 AA.
ID Y33C YEAST
AC P53280;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Hypothetical 128.8 kDa protein in PAS2-PRS5 intergenic region.
GN YGR134W.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RA van Dyck L., Skala J., de Wergifosse P., Purnelle B., Talia E.,
RA Nawrocki A., del Bino S., Goffeau A.;
RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: Z72919; CA97147.1; -.
DR PIR: S6443; S6443.
DR SCD: S0003366; CAF130.
DR GO: GO:0030014; C:CCR4-NOT complex; IPI.
DR GO: GO:0006357; P:regulation of transcription from pol II pro. .; IPI.
DR KW Hypothetical protein.
SQ SEQUENCE 1122 AA; 128779 MW; BF3542816CC78490 CRC64;

Query Match 37.0%; Score 44; DB 1; Length 1122;
Best Local Similarity 46.7%; Pred. No. 68;
Matches 7; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

```

```

Y3 1 ANNHVSIVEAMSDND 15
Db 110 SSVASYLSWKDND 124

RESULT 13
YPO4 CAEEL STANDARD; PRT; 1131 AA.
ID YPO4 CAEEL
AC P53015;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE Hypothetical 127.2 kDa protein F31E3.4 in chromosome III.
GN F31E3.4.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RA Favello A.;
RL Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: TO YEAST PAN2 AND TO S. POMBE SPAC22G7.04.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----

```

```

DR EMBL: U40935; AAA81687.1; -.
DR PIR: T16217; T16217.
DR Wormpep: F31E3.4; CE01269.
DR InterPro: IPR006055; Exonuclease.
DR Pfam: PF00929; Exonuclease; 1.
DR SMART: SM00479; EXOIII; 1.
DR Hypothetical protein.
SQ SEQUENCE 1131 AA; 127174 MW; D5544AB0C309E7DE CRC64;

Query Match 37.0%; Score 44; DB 1; Length 1131;
Best Local Similarity 36.8%; Pred. No. 69;
Matches 7; Conservative 5; Mismatches 7; Indels 0; Gaps 0;

QY 3 NHYSIVEAMSDNDTPYLHD 21
Db 315 NHGIVNVFADRDPOVNE 333

RESULT 14
PGCV_MOUSE STANDARD; PRT; 3358 AA.
ID PGCV_MOUSE STANDARD; PRT; 3358 AA.
AC Q62059; Q62058; Q9CTU0;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Versican core protein precursor (large fibroblast proteoglycan)
DE (Chondroitin sulfate proteoglycan core protein 2) (PG-M).
GN CSFG2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RC SEQUENCE FROM N.A. (ISOFORMS V0; V1 AND V2).
RP STRAIN=C57BL/6;
RX MEDLINE=95181355; PubMed=7876137;
RA Zako M., Shimomura T., Ujita M., Ito K., Kimata K.;
RT "Expression of Pg-M(V3), an alternatively spliced form of Pg-M
RT without a chondroitin sulfate attachment in region in mouse and human
RT tissues."
RT J. Biol. Chem. 270:3914-3918(1995).
RN [2]
RC SEQUENCE FROM N.A. (ISOFORM V3).
RP STRAIN=C57BL/6;
RX MEDLINE=95181355; PubMed=7876137;
RA Zako M., Shimomura T., Ujita M., Ito K., Kimata K.;
RT "Expression of Pg-M(V3), an alternatively spliced form of Pg-M
RT without a chondroitin sulfate attachment in region in mouse and human
RT tissues."
RT J. Biol. Chem. 270:3914-3918(1995).
RN [3]
RC SEQUENCE OF 1-1692 FROM N.A. (ISOFORM V1).
RP STRAIN=C57BL/6J; TISSUE=Skin;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Aizawa T., Hara A., Fukunishi Y., Komuro H., Adachi T., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamana I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nakai T., Pesole G., Quackenbush J.,
RA Schriml L.M., Stabili F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Blake J., Offield T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Brownstein M.T., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gueffroy S., Hill D., Hornum M., Hume D.A., Kamya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima U., Mazzarelli U., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyono O.K., Wang K.H., Welter C., Whitaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection."
RL Nature 409:685-690(2001).

```

```

RN [4]
RP INTERACTION WITH FBIN1.
RX PubMed=10400671;
RA Asberg A., Adam S., Kostka G., Timpel R., Heinegaard D.;
RT "Fibulin-1 is a ligand for the C-type lectin domains of aggrecan and
RT versican."
RT J. Biol. Chem. 274:20444-20449(1999).
RN [5]
RC FUNCTION: May play a role in intercellular signaling and in
RC connecting cells with the extracellular matrix. May take part in
CC the regulation of cell motility, growth and differentiation. Binds
CC hyaluronic acid.
CC SUBUNIT: Interacts with FBIN1.
CC -1- SUBCELLULAR LOCATION: Secreted; extracellular matrix.
CC -1- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=4;
CC Comment=Additional isoforms seem to exist;
CC Name=V0;
CC IsoId=Q62059-1; Sequence=Displayed;
CC Name=V1;
CC IsoId=Q62059-2; Sequence=VSP_003087; VSP_003088;
CC Name=V2;
CC IsoId=Q62059-3; Sequence=VSP_003089;
CC Name=V3;
CC IsoId=Q62059-4; Sequence=VSP_003087; VSP_003090;
CC -1- TISSUE SPECIFICITY: V2 is found only in brain.
CC -1- DEVELOPMENTAL STAGE: Disappears after the cartilage development.
CC -1- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.
CC -1- SIMILARITY: Contains 2 EGF-like domains.
CC -1- SIMILARITY: Contains 2 EGF-like domains.
CC -1- SIMILARITY: Contains 1 C-type lectin family domain.
CC -1- SIMILARITY: Contains 1 Sushi (SCR) domain.
CC -1- SIMILARITY: BELONGS TO THE AGGRECAN/VERSICAN PROTEOGLYCAN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.ebi.ac.uk/announcements
CC or send an email to license@ebi.ac.uk).
CC -----
DR EMBL: D12633; BAA03796.1; -.
DR EMBL: D28599; -; NOT ANNOTATED_CDS.
DR EMBL: D32040; BAA06802.1; -.
DR EMBL: AK014525; BAB29411.1; -.
DR HSSP: P01132; IEPG.
DR MGD: MGI:102889; Cspg2.
DR InterPro: IPR000152; Aex_hydroxyl.
DR InterPro: IPR000742; EGF_2.
DR InterPro: IPR001881; EGF_Ca.
DR InterPro: IPR006209; EGF_like.
DR InterPro: IPR007110; I9-like.
DR InterPro: IPR003599; I9-like.
DR InterPro: IPR003006; I9_MHC.
DR InterPro: IPR001304; Lectin_C.
DR InterPro: IPR000538; Link.
DR InterPro: IPR000436; Sushi_SCR_CCP.
DR Pfam: PF00008; EGF_2.
DR Pfam: PF00047; I9_1.
DR Pfam: PF00059; lectin_c_1.
DR Pfam: PF00084; sushi_1.
DR Pfam: PF00193; Xlink_2.
DR PRINTS: PR01265; LINKMODULE.
DR ProDom: PD000918; Link; 2.
DR SMART: SM00032; CCP; 1.
DR SMART: SM00034; CLEC7; 1.
DR SMART: SM00179; EGF_CA; 1.
DR SMART: SM00409; I9; 1.
DR SMART: SM00445; LINK; 2.
DR PROSITE: PS00010; ASX_HYDROXYL; 1.
DR PROSITE: PS00615; C_TYPE_LLECTIN_1; 1.
DR PROSITE: PS00041; C_TYPE_LLECTIN_2; 1.
DR PROSITE: PS00022; EGF_1; 2.

```

DR PROSITE; PS01186; EGF 2; 1.
 DR PROSITE; PS01187; EGF CA; 1.
 DR PROSITE; PS00835; IG Like; 1.
 DR PROSITE; PS01241; LINK; 2.
 KM Signal; Repeat; EGF-like domain; Calcium; Immunoglobulin domain;
 KM Hyaluronic acid; Alternative splicing.
 FT SIGNAL 1 20
 FT CHAIN 21 3358
 FT DOMAIN 21 146
 FT DOMAIN 167 244
 FT DOMAIN 265 346
 FT DOMAIN 348 1308
 FT DOMAIN 1309 3052
 FT DOMAIN 3052 3088
 FT DOMAIN 3090 3126
 FT DOMAIN 3139 3253
 FT DOMAIN 3258 3316
 FT DISULFID 44 130
 FT DISULFID 172 243
 FT DISULFID 196 217
 FT DISULFID 270 333
 FT DISULFID 294 315
 FT DISULFID 3061 3076
 FT DISULFID 3078 3087
 FT DISULFID 3094 3105
 FT DISULFID 3099 3114
 FT DISULFID 3116 3125
 FT DISULFID 3132 3143
 FT DISULFID 3160 3252
 FT DISULFID 3228 3244
 FT DISULFID 3259 3302
 FT DISULFID 3288 3315
 FT CARBOHYD 57 57
 FT CARBOHYD 330 330
 FT CARBOHYD 351 351
 FT CARBOHYD 441 441
 FT CARBOHYD 807 807
 FT CARBOHYD 914 914
 FT CARBOHYD 951 951
 FT CARBOHYD 1305 1305
 FT CARBOHYD 1372 1372
 FT CARBOHYD 1679 1679
 FT CARBOHYD 2054 2054
 FT CARBOHYD 2244 2244
 FT CARBOHYD 2362 2362
 FT CARBOHYD 2627 2627
 FT CARBOHYD 3030 3030
 FT CARBOHYD 3332 3332
 FT CARBOHYD 3342 3342
 FT CARBOHYD 348 348
 FT VARSPIC 349 1308
 FT VARSPIC 1309 3052
 FT VARSPIC 349 3052
 FT CONFLICT 126 126
 FT CONFLICT 348 348
 FT CONFLICT 1658 1658
 FT CONFLICT 1674 1680
 FT SEQUENCE 3358 AA; 366938 MW; 07180026BC0762D CRC64;
 Query Match 37.0%; Score 44; DB 1; Length 3358;
 Best Local Similarity 50.0%; Freq. No. 2.6e+02;
 Matches 8; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

RESULT 15
 ID G6PD_ECOLI STANDARD: PRT: 491 AA.
 AC P22992; P78069; Q60134; Q60139;
 DT 01-AUG-1991 (Rel. 19, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 26-FEB-2003 (Rel. 41, Last annotation update)
 DE Glucose-6-phosphate 1-dehydrogenase (EC 1.1.1.49) (G6PD).
 GN ZWF OR B1852 OR C2265.
 OS Escherichia coli, and
 OS Escherichia coli O6.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Escherichia.
 OC NCBI_TaxId=562, 217992;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=K12;
 RX MEDLINE=91123224; PubMed=1704005;
 RT Rowley D.L., Wolf R.E. Jr.;
 RL "Molecular characterization of the Escherichia coli K-12 zwf gene encoding glucose 6-phosphate dehydrogenase.";
 RN J. Bacteriol. 173:968-977(1991).
 [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=K12 / MG1655;
 RX MEDLINE=9742617; PubMed=9278503;
 RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
 RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
 RA Gregor U., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
 RA Mau B., Shao Y.;
 RT "The complete genome sequence of Escherichia coli K-12.";
 RL Science 277:1234-1238(1997).
 [3]
 RP SEQUENCE FROM N.A.
 RC MEDLINE=97251358; PubMed=9097040;
 RA Itoh T., Aiba H., Baba T., Fujita K., Hayashi K., Inada T.,
 RA Isono K., Kasai H., Kimura S., Kitakawa M., Kitakawa M.,
 RA Makino K., Maki T., Mizobuchi K., Mori H., Mori T., Motomura K.,
 RA Nakade S., Nakamura Y., Nishimoto H., Nishio Y., Oshita H.,
 RA Saito N., Samped G., Seki Y., Sivasubramanian S., Tagami H.,
 RA Takekida U., Takemoto K., Mada C., Yamamoto Y., Horinouchi T.;
 RT "A 460-kb DNA sequence of the Escherichia coli K-12 genome corresponding to the 40.1-50.0 min region on the linkage map.";
 RL DNA Res. 3:379-392(1996).
 [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN=06:H1 / CF7073 / ATCC 700928;
 RX MEDLINE=2288234; PubMed=12471157;
 RA Welch R.A., Burland V., Plunkett G. III, Redford P., Roesch P.,
 RA Rasko D., Buckles E.L., Liu S.-R., Boutin A., Hackett J., Stroud D.,
 RA Mayhew G.F., Rose D.J., Zhou S., Schwartz D.C., Perna N.T.,
 RA Mobley H.L.T., Donnenberg M.S., Blattner F.R.;
 RT "Extensive mosaic structure revealed by the complete genome sequence of uropathogenic Escherichia coli.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:17020-17024(2002).
 [5]
 RP SEQUENCE OF 72-368 FROM N.A.
 RC STRAIN=Various ECOLI strains;
 RX MEDLINE=95064015; PubMed=7973728;
 RA Gutman D.S., Dykhuizen D.E.;
 RT "Clonal divergence in Escherichia coli as a result of recombination, not mutation.";
 RL Science 266:1380-1383(1994).
 [6]
 RP SEQUENCE OF 321-491 FROM N.A.
 RC STRAIN=K12;
 RX MEDLINE=93345818; PubMed=8344525;
 RA Carter A.T., Pearson B.M., Dickinson J.R., Larcaschire W.E.;
 RT "Sequence of the Escherichia coli K-12 edd and eda genes of the Brunner-Doudoroff pathway.";

```

Gene 130:155-156(1993).
CC -1- CATALYTIC ACTIVITY: D-glucose 6-phosphate + NADP(+) = D-glucono-
CC 1,5-lactone 6-phosphate + NADPH.
CC -1- PATHWAY: Pentose phosphate pathway, first step.
CC -1- SIMILARITY: BELONGS TO THE GLUCOSE-6-PHOSPHATE DEHYDROGENASE
CC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC
DR EMBL, M55005, AAA24775.1, -.
DR EMBL, AE000279, AAC74922.1, -.
DR EMBL, D90828, EAA15660.1, -.
DR EMBL, AE016761, AAN80722.1, -.
DR EMBL, U13783, AAA57018.1, -.
DR EMBL, U13784, AAA57019.1, -.
DR EMBL, U13785, AAA57020.1, -.
DR EMBL, U13786, AAA57021.1, -.
DR EMBL, U13787, AAA57022.1, -.
DR EMBL, U13788, AAA57023.1, -.
DR EMBL, U13789, AAA57024.1, -.
DR EMBL, U13790, AAA57025.1, -.
DR EMBL, U13791, AAA57026.1, -.
DR EMBL, U13792, AAA57027.1, -.
DR EMBL, U13793, AAA57028.1, -.
DR EMBL, U13794, AAA57029.1, -.
DR EMBL, X63694, CAA45220.1, -.
DR PIR, D64947, D64947.
DR HSSP, P11411, IDPG.
DR SWISS-2DPAGE, P22992, COLI.
DR EC02DBASE, F048.8, 6TH EDITION.
DR EcoGene, EG1121, zwf.
DR InterPro, IPR001282, G6PD.
DR Pfam, PF00479, G6PD. 1.
DR Pfam, PF02781, G6PD_C. 1.
DR PRINTS, PR000079, G6PDHDSGNASE.
DR PRODOM, PD001129, G6PD, 1.
DR TIGRPFAMs, TIGR00871, zwf, 1.
DR PROSITE, PS00069, G6P_DEHYDROGENASE, 1.
DR Oxidoreductase, NADP, Glucose metabolism, Complete proteome.
KW ACT SITE 181
FT VARIANT 100 100
FT CONFLICT 268 293
FT
SQ SEQUENCE 491 AA, 55704 MW, 263F07D298EAFCD3 CRC64;
Query Match 36.6%; Score 43.5; DB 1; Length 491;
Best local Similarity 69.2%; Pred. No. 30;
Matches 9; Conservative 1; Mismatches 2; Indels 1; Gaps 1.

QY 6 SIVEANS-DNDTP 17
|||
|||
|||
|||
Db 451 SITEAMAMDNDAP 463

Search completed: November 13, 2003, 09:45:31
Job time : 8.56398 secs

```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using SW model

Run on: November 13, 2003, 09:31:40 / Search time 31.6493 Seconds
(without alignments)
171.224 Million cell updates/sec

Title: US-09-290-049a-12
Perfect score: 119
Sequence: 1 ANNHVSIVEAWSNDPTPLHD 21

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 830525 seqs, 258052604 residues
Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database:

SPREMBL_23:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phage:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_virus:*
- 16: sp_bacteriaph:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	length	DB ID	Description
1	119	100.0	1590	2	Q59983 streptococc
2	119	100.0	1590	2	Q55263 streptococc
3	76	63.9	1016	2	Q91CJ7 leuconostoc
4	72	60.5	2835	2	Q8G9Q2 leuconostoc
5	68	57.1	1477	2	Q91466 leuconostoc
6	68	57.1	1508	2	Q9E2H5 leuconostoc
7	68	57.1	1508	2	Q52224 leuconostoc
8	68	57.1	1575	2	Q91CH3 streptococc
9	68	57.1	1577	2	Q54178 streptococc
10	67	56.3	1338	2	Q9WXJ4 streptococc
11	65	54.6	1527	2	Q9ZAR4 leuconostoc
12	65	54.6	1527	2	Q8KRE1 leuconostoc
13	61	51.3	1449	2	Q68542 streptococc
14	61	51.3	1449	2	Q55264 streptococc
15	61	51.3	1554	2	Q8KZL5 streptococc
16	60	50.4	1512	2	Q9WXJ5 streptococc

17	57	47.9	1577	2	Q55265 streptococc
18	56	47.1	1518	2	Q06600 streptococc
19	55	46.2	1599	2	Q00599 streptococc
20	55	46.2	2057	2	Q9RE05 leuconostoc
21	53	44.5	175	10	Q91MJ8 arabidopsis
22	53	44.5	195	10	Q8LDM8 arabidopsis
23	51	42.9	1195	2	Q06459 thermus the
24	50	42.0	601	5	Q9VOR4 drosophila
25	50	42.0	630	5	Q81FA3 drosophila
26	49	41.2	524	16	Q9KSI8 drosophila
27	48.5	40.8	282	10	Q8RWV3 drosophila
28	48	40.3	184	16	Q8XME4 drosophila
29	48	40.3	380	3	Q8N773 drosophila
30	47	39.5	193	2	Q9K445 drosophila
31	46.5	39.1	212	2	Q9K167 drosophila
32	46.5	39.1	1604	4	Q8NPA0 drosophila
33	46.5	39.1	1604	5	Q8S589 drosophila
34	46	38.7	169	5	Q9VTP7 drosophila
35	46	38.7	271	11	Q9QZEE drosophila
36	46	38.7	283	5	Q8MSF4 drosophila
37	46	38.7	283	5	Q8Y2Q6 drosophila
38	46	38.7	393	16	Q8Y2Q6 drosophila
39	46	38.7	490	16	Q06579 mycobacteri
40	46	38.7	601	16	Q05566 mycobacteri
41	46	38.7	810	5	Q8T3J2 drosophila
42	46	38.7	811	5	Q9VK54 drosophila
43	46	38.7	880	5	P91643 drosophila
44	45.5	38.2	418	17	Q97U15 sulfolobus
45	45	37.8	119	2	Q93RH1 streptococc

ALIGNMENTS

RESULT 1

Q59983 PRELIMINARY; PRT; 1590 AA.

AC Q59983; 01-NOV-1996 (TREMUR1.01, Created)
 DT 01-NOV-1996 (TREMUR1.01, Last sequence update)
 DT 01-OCT-2002 (TREMUR1.22, Last annotation update)
 DE Glucosyltransferase-I precursor (EC 2.4.1.5).
 GN GTFI.
 OS Streptococcus sobrinus.
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 CC Streptococcus.
 OX NCBI_TaxID=1310;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=QW176;
 RA MEDLINE=94146405; PubMed=8312602;
 RX Sato S., Inoue M., Handa N., Aizawa Y., Isobe Y., Karayama T.;
 RT "DNA sequence of the glucosyltransferase gene of serotype d
 Streptococcus sobrinus.";
 RL DNA Seq. 4:19-27(1993).
 DR EMBL, D13858; BAA02976.1; -;
 DR InterPro; IPR002479; CW_binding.
 DR InterPro; IPR003318; Glyco_hydro_70.
 DR Pfam; PF01473; CW_binding_1; 16.
 DR Pfam; PF02324; Glyco_hydro_70; 1.
 KW Glycosyltransferase; Signal; Transferase.
 FT SIGNAL 1
 FT CHAIN 39 1590 GLUCOSYLTRANSFERASE-I.
 FT SEQUENCE 1590 AA; 17595 MW; C3C83A57CF3C280E CRC64;

Query Match 100.0%; Score 119; DB 2; Length 1590;
 Best Local Similarity 100.0%; Pred. No. 1,2e-09;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 ANNHVSIVEAWSNDPTPLHD 21
 DB 477 ANNHVSIVEAWSNDPTPLHD 497

RESULT 2

Q55263 PRELIMINARY; PRT; 1590 AA.
 AC Q55263;
 DT 01-NOV-1996 (Tremblrel. 01, Created)
 DT 01-JAN-1998 (Tremblrel. 05, Last sequence update)
 DT 01-OCT-2002 (Tremblrel. 22, Last annotation update)
 DE GTF-1.
 GN GLUCOSYLTRANSFERASE.
 OS Streptococcus sobrinus.
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 CX NCBI_TaxID=1310;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 33478;
 RA Sato S.;
 RT "DNA and amino-acid sequences of water-insoluble-glucan synthetase produced from Streptococcus sobrinus ATCC 33478."
 RL Ann. Kagoshima Univ. Dental School 16:23-29(1996).
 DR EMBL; D63570; BAA09792.1; -.
 DR InterPro; IPR002479; CW binding.
 DR InterPro; IPR003318; Glyco_hydro_70.
 DR Pfam; PF01473; CW binding_1; 15.
 DR Pfam; PF02324; Glyco_hydro_70; 1.
 DR SEQUENCE 1590 AA; 176058 MW; 9DF7A3F2C6E4FD43 CRC64;

Query Match 100.0%; Score 119; DB 2; Length 1590;
 Best Local Similarity 100.0%; Pred. No. 1.2e-09;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ANNHVSIVEAWSNDNTPYLHD 21
 ID 477 ANNHVSIVEAWSNDNTPYLHD 497

RESULT 3
 Q91CJ7 PRELIMINARY; PRT; 1016 AA.
 AC Q91CJ7;
 DT 01-OCT-2000 (Tremblrel. 15, Created)
 DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
 DT 01-OCT-2002 (Tremblrel. 22, Last annotation update)
 DE Dextranucrase.
 GN DSRT.
 OS Leuconostoc mesenteroides.
 OC Bacteria; Firmicutes; Lactobacillales; Leuconostoc.
 CX NCBI_TaxID=1245;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=NRRL B-512F;
 RX MEDLINE=20169623; PubMed=10705445; Kobayashi M.;
 RA Funane K., Mizuno K., Takahara H.;
 RT "Gene encoding a dextranucrase-like protein in Leuconostoc mesenteroides NRRL B-512F."
 RL EMBL; A8020020; BAA90527.1; -.
 DR HSBP; P06278; IVTS.
 DR InterPro; IPR003318; Glyco_hydro_70.
 DR Pfam; PF02324; Glyco_hydro_70; 1.
 DR SEQUENCE 1016 AA; 110344 MW; 8896EFDE13CCCB47 CRC64;

Query Match 63.9%; Score 76; DB 2; Length 1016;
 Best Local Similarity 61.9%; Pred. No. 0.0042;
 Matches 13; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

Qy 1 ANNHVSIVEAWSNDNTPYLHD 21
 ID 548 ANNHVSIVEAWSNDNTPYLHD 568

RESULT 4

Q8G9Q2 PRELIMINARY; PRT; 2835 AA.
 AC Q8G9Q2;
 DT 01-MAR-2003 (Tremblrel. 23, Created)
 DT 01-MAR-2003 (Tremblrel. 23, Last sequence update)
 DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
 DE Dextranucrase (EC 2.4.1.5) (Fragment).
 GN DSRE.
 OS Leuconostoc mesenteroides.
 OC Bacteria; Firmicutes; Lactobacillales; Leuconostoc.
 CX NCBI_TaxID=1245;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=2231661; PubMed=12270834;
 RA Bozonnet S., Dols-Lafargue M., Fabre E., Pizzut S., Renaud-Simeon M.,
 RA Mondan P., Willemot R.M.;
 RT "Molecular characterization of DSR-E, an alpha-1,2 linkage synthesizing dextranucrase with two catalytic domains."
 RL J. Bacteriol. 184:5753-5761(2002).
 DR EMBL; AJ430204; CAD22883.1; -.
 DR Transferase; Glycosyltransferase.
 KW NON TER
 FT 1
 SQ SEQUENCE 2835 AA; 313264 MW; D03262CDD735399D CRC64;

Query Match 60.5%; Score 72; DB 2; Length 2835;
 Best Local Similarity 57.9%; Pred. No. 0.058;
 Matches 11; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

Qy 1 ANNHVSIVEAWSNDNTPYL 19
 ID 557 ANNHVSIVEAWSNDNTPYL 575

RESULT 5
 Q91A66 PRELIMINARY; PRT; 1477 AA.
 AC Q91A66;
 DT 01-OCT-2000 (Tremblrel. 15, Created)
 DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
 DT 01-OCT-2002 (Tremblrel. 22, Last annotation update)
 DE Dextranucrase (EC 2.4.1.5).
 GN DSRC.
 OS Leuconostoc mesenteroides.
 OC Bacteria; Firmicutes; Lactobacillales; Leuconostoc.
 CX NCBI_TaxID=1245;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=NRRL B-1355;
 RA Arguello-Morales M.A., Renaud-Simeon M., Pizzut S., Sarcabal P.,
 RA Willemot R.M., Monsan P.;
 RT "Sequence analysis of the gene encoding alternansucrase, a sucrose RT glucosyltransferase from Leuconostoc mesenteroides NRRL B-1355."
 RL Submitted (OCT-1999) to the EMBL/Genbank/DBS databases.
 DR EMBL; AJ250172; CAB76565.1; -.
 DR InterPro; IPR002479; CW binding.
 DR InterPro; IPR003318; Glyco_hydro_70.
 DR Pfam; PF01473; CW binding_1; 14.
 DR Pfam; PF02324; Glyco_hydro_70; 1.
 DR Glycosyltransferase; Transferase.
 KW Glycosyltransferase; Transferase.
 SQ SEQUENCE 1477 AA; 164887 MW; 66F5710DEDFCB831 CRC64;

Query Match 57.1%; Score 68; DB 2; Length 1477;
 Best Local Similarity 57.1%; Pred. No. 0.12;
 Matches 12; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

Qy 1 ANNHVSIVEAWSNDNTPYLHD 21
 ID 532 ANNHVSIVEAWSNDNTPYLHD 552

RESULT 6
 Q9EZHS PRELIMINARY; PRT; 1508 AA.
 ID Q9EZHS

```

AC Q95ZH5;
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
DE Dextranucrase DsrB742.
GN DSRB742.
OS Leuconostoc mesenteroides.
OC Bacteria; Firmicutes; Lactobacillales; Leuconostoc.
OX NCBI_TaxID=1245;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=B-742CB;
RA Kim H.-S., Kim D., Ryu H.-J., Rohyt J.F.;
RT "Leuconostoc mesenteroides B-742CB, a dextranucrase gene.";
RL Submitted (Aug-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF294469; AAG38021.1; -
DR InterPro; IPR002479; CW_binding.
DR InterPro; IPR003318; Glyco_hydro_70.
DR Pfam; PF01473; CW_binding_1; 14.
DR Pfam; PF02324; Glyco_hydro_70; 1.
SQ SEQUENCE 1508 AA; 168542 MW; E2FCFA0F87A8AF3A CRC64;

Query Match 57.1%; Score 68; DB 2; Length 1508;
Best Local Similarity 57.1%; Pred. No. 0.12;
Matches 12; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 1 ANNHVSIVEAWSNDPTPLMD 21
Db 563 ANOHLSTLEDSHNDPEYKD 583

RESULT 7
ID 052224 PRELIMINARY; PRT; 1508 AA.
AC 052224;
DT 01-JUN-1998 (TREMBlrel. 06, Created)
DT 01-JUN-1998 (TREMBlrel. 06, Last sequence update)
DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
DE Glucosyltransferase (EC 2.4.1.5).
GN DSRB.
OS Leuconostoc mesenteroides.
OC Bacteria; Firmicutes; Lactobacillales; Leuconostoc.
OX NCBI_TaxID=1245;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NRRL B-1299;
RA Monchois V., Renaud-Simeon M., Monsan P., Willemot R.M.;
RT "Cloning and sequencing of a gene coding for an extracellular
RT dextranucrase (DSRB) from Leuconostoc mesenteroides NRRL B-1299
RT synthesizing only a a(1-6) glucan.";
RL EMBL Microbiol. Lett. 0:0-0(1998).
DR EMBL; AF030129; AAB95453.1; -
DR InterPro; IPR002479; CW_binding.
DR InterPro; IPR003318; Glyco_hydro_70.
DR Pfam; PF01473; CW_binding_1; 14.
DR Pfam; PF02324; Glyco_hydro_70; 1.
DR Pfam; PF02324; Glyco_hydro_70; 1.
KW Glycosyltransferase; Transferase.
SQ SEQUENCE 1508 AA; 168511 MW; E70CECB57A70D1F0 CRC64;

Query Match 57.1%; Score 68; DB 2; Length 1508;
Best Local Similarity 57.1%; Pred. No. 0.12;
Matches 12; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 1 ANNHVSIVEAWSNDPTPLMD 21
Db 563 ANOHLSTLEDSHNDPEYKD 583

RESULT 8
ID 09LCH3 PRELIMINARY; PRT; 1575 AA.
AC 09LCH3;
DT 01-OCT-2000 (TREMBlrel. 15, Created)

```

```

DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-OCT-2000 (TREMBlrel. 22, Last annotation update)
DE Glucosyltransferase.
GN GTRF.
OS Streptococcus oralis.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1303;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC10557;
RX MEDLINE=20231779; PubMed=10768934;
RA Fujiwara T., Hoshino T., Oshima T., Sobue S., Hamada S.;
RT "Purification, characterization, and molecular analysis of the gene
RT encoding glucosyltransferase from Streptococcus oralis.";
RL Infect. Immun. 68:2475-2483(2000).
DR EMBL; AB025228; BA95201.1; -
DR InterPro; IPR002479; CW_binding.
DR InterPro; IPR003318; Glyco_hydro_70.
DR Pfam; PF01473; CW_binding_1; 17.
DR Pfam; PF02324; Glyco_hydro_70; 1.
KW Transferase.
SQ SEQUENCE 1575 AA; 176792 MW; 772A26E4D7C2E543 CRC64;

Query Match 57.1%; Score 68; DB 2; Length 1575;
Best Local Similarity 61.9%; Pred. No. 0.13;
Matches 13; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 1 ANNHVSIVEAWSNDPTPLMD 21
Db 546 AIRHLSLLEAWSNDPEYKD 566

RESULT 9
ID 054178 PRELIMINARY; PRT; 1577 AA.
AC 054178; Q54247;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
DE Glucosyltransferase.
GN GTRF.
OS Streptococcus gordonii Challis.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=29390;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CHALLIS;
RX MEDLINE=96157084; PubMed=8586195;
RA Vickerman M.M., Sulavik M.C., Clewell D.B.;
RT "Molecular analysis of Streptococcus gordonii glucosyltransferase
RT phase variants.";
RL Dev. Biol. Stand. 85:309-314(1995).
RN [2]
RP SEQUENCE OF 1-96 FROM N.A.
RC STRAIN=CHALLIS;
RX MEDLINE=92276337; PubMed=1534326;
RA Sulavik M.C., Tardif G., Clewell D.B.;
RT "Identification of a gene, rgg, which regulates expression of
RT glucosyltransferase and influences the spp phenotype of Streptococcus
RT gordonii Challis.";
RL J. Bacteriol. 174:3577-3586(1992).
DR EMBL; U12643; AAC43483.1; -
DR EMBL; M89776; AAB26969.1; -
DR InterPro; IPR002479; CW_binding.
DR InterPro; IPR003318; Glyco_hydro_70.
DR Pfam; PF01473; CW_binding_1; 18.
DR Pfam; PF02324; Glyco_hydro_70; 1.
KW Transferase.
SQ SEQUENCE 1577 AA; 177805 MW; 5AF0328DC5E08D18 CRC64;

Query Match 57.1%; Score 68; DB 2; Length 1577;

```

Best Local Similarity 61.9%; Pred. No. 0.13;
Matches 13; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 1 ANNHVSIVEAWSNDPPLYHD 21
Db 548 ALKHLSTLEAWSNDPPLYHD 568

RESULT 10

Q9WXJ4 ID Q9WXJ4 PRELIMINARY; PRT; 1338 AA.
AC Q9WXJ4;
DT 01-NOV-1999 (TREMBlrel. 12, Created)
DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
DE GTF-S.
GN GTF-S.
OS Streptococcus criceti.
OG Plasmid PAM1.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1333;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HS-6;
RA Inoue M., Fukui K., Miyagi A.;
RT "S. citreus glucosyltransferase (glfs and gtf) genes."
RL Submitted (MAR-1999) to the EMBL/Genbank/DBJ databases.
DR EMBL; AB026123; BAA7236.1; -
DR InterPro; IPR002479; CW binding.
DR InterPro; IPR003318; Glyco_hydro_70.
DR Pfam; PF01473; CW binding_1; 10.
DR Pfam; PF02324; Glyco_hydro_70; 1.
KM Plasmid.
SQ SEQUENCE 1338 AA; 148558 MW; 0A90C8E10E1SD99B CRC64;

Query Match 56.3%; Score 67; DB 2; Length 1338;
Best Local Similarity 57.1%; Pred. No. 0.15;
Matches 12; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 1 ANNHVSIVEAWSNDPPLYHD 21
Db 437 AIDHLSLEAWSNDPPLYHD 457

RESULT 11

Q9ZAR4 ID Q9ZAR4 PRELIMINARY; PRT; 1527 AA.
AC Q9ZAR4;
DT 01-MAY-1999 (TREMBlrel. 10, Created)
DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)
DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
DE Dextranucrase.
GN DEX.
OS Leuconostoc mesenteroides.
OC Bacteria; Firmicutes; Lactobacillales; Leuconostoc.
OX NCBI_TaxID=1245;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NRRL B-512-F;
RA Bhattacharjee R., Singh D.K.S.;
RT "Cloning and Molecular Characterization of Dextranucrase Gene from
Leuconostoc mesenteroides NRRL B-512-F."
RL Submitted (DEC-1996) to the EMBL/Genbank/DBJ databases.
DR EMBL; U81374; AAD10952.1; -
DR InterPro; IPR002479; CW binding.
DR InterPro; IPR003318; Glyco_hydro_70.
DR Pfam; PF01473; CW binding_1; 16.
DR Pfam; PF02324; Glyco_hydro_70; 1.
SQ SEQUENCE 1527 AA; 169709 MW; 1DEFA237C743398 CRC64;

Query Match 54.6%; Score 65; DB 2; Length 1527;
Best Local Similarity 57.1%; Pred. No. 0.36;

Matches 12; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 1 ANNHVSIVEAWSNDPPLYHD 21
Db 581 ANQHLSTLEAWSNDPPLYHD 601

RESULT 12

Q8KRE1 ID Q8KRE1 PRELIMINARY; PRT; 1527 AA.
AC Q8KRE1;
DT 01-OCT-2002 (TREMBlrel. 22, Created)
DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Dextranucrase DsrD (EC 2.4.1.5).
GN DSRD.
OS Leuconostoc mesenteroides.
OC Bacteria; Firmicutes; Lactobacillales; Leuconostoc.
OX NCBI_TaxID=1245;
RN [1]
RP SEQUENCE FROM N.A.
RA Neubauer H., Bauche A., Mollet B.;
RT "Isolation and characterization of the dextranucrase DsrD of
Leuconostoc mesenteroides Lcc4."
RL Submitted (JAN-2001) to the EMBL/Genbank/DBJ databases.
DR EMBL; AY017384; AAG6158.1; -
DR InterPro; IPR002479; CW binding.
DR InterPro; IPR003318; Glyco_hydro_70.
DR Pfam; PF01473; CW binding_1; 6.
DR Pfam; PF02324; Glyco_hydro_70; 1.
DR Transferrase; Glycosyltransferase.
KM Transferrase.
SQ SEQUENCE 1527 AA; 169835 MW; F9D0DE220BD93668 CRC64;

Query Match 54.6%; Score 65; DB 2; Length 1527;
Best Local Similarity 57.1%; Pred. No. 0.36;
Matches 12; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 1 ANNHVSIVEAWSNDPPLYHD 21
Db 581 ANQHLSTLEAWSNDPPLYHD 601

RESULT 13

O68542 ID O68542 PRELIMINARY; PRT; 1449 AA.
AC O68542;
DT 01-AUG-1998 (TREMBlrel. 07, Created)
DT 01-AUG-1998 (TREMBlrel. 07, Last sequence update)
DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
DE Glucosyltransferase N (Fragment).
GN GTFN.
OS Streptococcus salivarius.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1304;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VLA477;
RA Jaffe R.I.;
RT "Streptococcus salivarius VLA477 gtfN."
RL Submitted (FEB-1998) to the EMBL/Genbank/DBJ databases.
DR EMBL; AF049609; AAC05156.1; -
DR InterPro; IPR002479; CW binding.
DR InterPro; IPR003318; Glyco_hydro_70.
DR Pfam; PF01473; CW binding_1; 8.
DR Pfam; PF02324; Glyco_hydro_70; 1.
KM Transferrase.
KW NON TER
PT NON TER
SQ SEQUENCE 1449 AA; 159895 MW; 0700FED748471BFB CRC64;

Query Match 51.3%; Score 61; DB 2; Length 1449;
Best Local Similarity 57.1%; Pred. No. 1.4;
Matches 12; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

OY 1 ANNHVSIVEAWSNDPTYLHD 21
 DB 536 AIKHLSILEAWSHNDAYTNEED 556

RESULT 14
 ID 055264 PRELIMINARY; PRT; 1449 AA.
 AC 055264;
 DT 01-NOV-1996 (TREMBlrel. 01, Created)
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
 DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
 DE Glucosyltransferase precursor.
 GN GTF.
 OS Streptococcus salivarius.
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 OX NCBI_TaxID=1304;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=95122197; PubMed=7822030;
 RA Simpson C.L., Giffard P.M., Jacques N.A.;
 RT "Streptococcus salivarius ATCC 25975 possesses at least two genes
 coding for primer-independent glucosyltransferases.";
 RL Infect. Immun. 63:609-621(1995).
 DR EMBL; L35495; AAC41412.1; -;
 DR InterPro; IPR002479; CW binding.
 DR InterPro; IPR003318; Glyco_hydro_70.
 DR Pfam; PF01473; CW_binding_1; 8.
 DR Pfam; PF02324; Glyco_hydro_70; 1.
 KM Signal; transferase.
 FT SIGNAL 1 35 POTENTIAL.
 FT CHAIN 36 1449 GLUCOSYLTRANSFERASE.
 SQ SEQUENCE 1449 AA; 159984 MW; D62F07306E86A46 CRC64;
 Query Match 51.3%; Score 61; DB 2; Length 1449;
 Best Local Similarity 57.1%; Pred. No. 1.4;
 Matches 12; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

OY 1 ANNHVSIVEAWSNDPTYLHD 21
 DB 536 AIKHLSILEAWSHNDAYTNEED 556

RESULT 15
 ID 08KZL5 PRELIMINARY; PRT; 1554 AA.
 AC 08KZL5;
 DT 01-OCT-2002 (TREMBlrel. 22, Created)
 DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)
 DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
 DE Glucosyltransferase.
 GN GTF.
 OS Streptococcus sobrinus.
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 OX NCBI_TaxID=1310;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21958684; PubMed=11960691;
 RA Hanada N., Fukushima K., Nomura Y., Sengoku H., Hayakawa M.,
 RA Mukasa H., Shiroza T., Abiko Y.;
 RT "Cloning and nucleotide sequence analysis of the Streptococcus
 RT sobrinus gtfu gene that produces a highly branched water-soluble
 RT glucan.";
 RL Biochim. Biophys. Acta 1570:75-79(2002).
 DR EMBL; AB089438; BAC07265.1; -;
 DR InterPro; IPR002479; CW_binding.
 DR InterPro; IPR003318; Glyco_hydro_70.
 DR Pfam; PF01473; CW_binding_1; 6.
 DR Pfam; PF02324; Glyco_hydro_70; 1.
 KM Transferase.

SQ SEQUENCE 1554 AA; 171676 MW; 6981BC1DAE24A73 CRC64;
 Query Match 51.3%; Score 61; DB 2; Length 1554;
 Best Local Similarity 61.1%; Pred. No. 1.6;
 Matches 11; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

OY 4 HVSIVEAWSNDPTYLHD 21
 DB 488 HISILEAWSINDQYNEED 505

Search completed: November 13, 2003, 09:44:04
 Job time : 32.6493 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 13, 2003, 09:31:40 ; Search time 13.6351 Seconds
(without alignments)
65.165 Million cell updates/sec

Title: US-09-290-049A-12
Perfect score: 119
Sequence: 1 ANNHVSIVFAMSDNDTPYLHD 21

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/ptodata/2/1aa/5A.COMB.pep.*
2: /cgn2_6/ptodata/2/1aa/5B.COMB.pep.*
3: /cgn2_6/ptodata/2/1aa/6A.COMB.pep.*
4: /cgn2_6/ptodata/2/1aa/6B.COMB.pep.*
5: /cgn2_6/ptodata/2/1aa/PCPUS.COMB.pep.*
6: /cgn2_6/ptodata/2/1aa/backfile1.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	108	90.8	1475	3	US-09-007-999-2
2	108	90.8	1475	3	US-09-210-361-2
3	108	90.8	1475	4	US-09-740-274-2
4	99	83.2	1375	3	US-09-210-361-4
5	99	83.2	1375	4	US-09-740-274-4
6	74	62.2	1430	3	US-09-604-957-4
7	74	62.2	1430	3	US-09-008-172-2
8	74	62.2	1430	3	US-09-210-361-6
9	74	62.2	1430	3	US-09-740-274-6
10	65	54.6	523	4	US-09-604-957-5
11	57	47.9	1577	2	US-08-793-824-2
12	55	46.2	584	4	US-09-604-957-6
13	55	46.2	2057	4	US-09-499-203-2
14	48	40.3	535	4	US-09-604-957-7
15	48	40.3	1278	4	US-09-604-957-3
16	47	39.5	749	4	US-09-562-737-96
17	43.5	36.6	1092	4	US-09-275-608-3
18	43	36.1	151	4	US-09-198-452A-815
19	43	36.1	234	4	US-09-134-001C-4514
20	43	36.1	647	4	US-09-134-001C-5458
21	42	35.3	200	2	US-08-836-442-4
22	42	35.3	287	4	US-09-252-991A-31548
23	42	35.3	327	4	US-09-107-532A-6181
24	42	35.3	334	6	5290690-11
25	42	35.3	403	4	US-09-252-991A-27960
26	42	35.3	454	4	US-09-252-991A-28000
27	42	35.3	455	3	US-09-162-473-6

28	42	35.3	529	2	US-08-836-442-3	Sequence 3, Appli
29	42	35.3	1627	1	US-07-665-792E-9	Sequence 9, Appli
30	42	35.3	3898	3	US-08-750-717-2	Sequence 2, Appli
31	41.5	34.9	69	4	US-09-252-991A-17383	Sequence 17383, A
32	41.5	34.9	130	4	US-09-328-352-4952	Sequence 4952, Ap
33	41	34.5	88	4	US-09-252-991A-21480	Sequence 21480, A
34	41	34.5	259	1	US-08-277-231A-3	Sequence 3, Appli
35	41	34.5	259	2	US-08-473-750-6	Sequence 6, Appli
36	41	34.5	259	2	US-08-477-326-6	Sequence 3, Appli
37	41	34.5	2465	3	US-08-596-291-3	Sequence 3, Appli
38	41	34.5	2465	2	US-09-100-804-3	Sequence 12, Appli
39	41	34.5	2465	4	US-09-080-855-12	Sequence 12, Appli
40	41	34.5	2466	4	US-09-566-076-12	Sequence 2, Appli
41	41	34.5	2466	5	PCT-US94-09943-2	Sequence 46, Appli
42	41	34.5	2485	3	US-09-290-640-46	Sequence 2, Appli
43	41	34.5	3898	2	US-08-876-991-2	Sequence 2, Appli
44	41	34.5	3898	2	US-09-059-853-2	Sequence 2, Appli
45	40.5	34.0	214	3	US-09-214-278-1	Sequence 1, Appli

ALIGNMENTS

```
RESULT 1
US-09-007-999-2
; Sequence 2, Application US/09007999
; Patent No. 6087559
; GENERAL INFORMATION:
; APPLICANT: Nichols, Scott E.
; TITLE OF INVENTION: Substitutes for Modified Starch and
; FILE REFERENCE: Latexes in Paper Manufacture
; CURRENT APPLICATION NUMBER: US/09/007,999
; CURRENT FILING DATE: 1998-01-16
; EARLIER APPLICATION NUMBER: 08/478,704
; EARLIER FILING DATE: 1995-06-07
; NUMBER OF SEQ. ID NOS: 2
; SOFTWARE: FASTSEQ for Windows Version 3.0
; SEQ. ID NO 2
; LENGTH: 1475
; TYPE: PRT
; ORGANISM: Streptococcus mutans
US-09-007-999-2

Query Match      90.8%; Score 108; DB 3; Length 1475;
Best Local Similarity 85.7%; Pred. No. 1.3e-08;
Matches 18; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Cy      1 ANNHVSIVFAMSDNDTPYLHD 21
Db      481 ANNHVSIVFAMSDNDTPYLHD 501

RESULT 2
US-09-210-361-2
; Sequence 2, Application US/09210361
; Patent No. 6284479
; GENERAL INFORMATION:
; APPLICANT: Nichols, Scott E.
; TITLE OF INVENTION: Substitutes for Modified Starches and
; FILE REFERENCE: Latexes in Paper Manufacture
; CURRENT APPLICATION NUMBER: US/09/210,361
; CURRENT FILING DATE: 1998-12-11
; EARLIER APPLICATION NUMBER: 09/007,999
; EARLIER FILING DATE: 1998-01-16
; EARLIER APPLICATION NUMBER: 08/478,704
; EARLIER FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 09/009,620
; EARLIER FILING DATE: 1998-01-20
; EARLIER APPLICATION NUMBER: 08/485,243
; EARLIER FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 09/008,172
```

```

; EARLIER FILING DATE: 1998-01-16
; EARLIER APPLICATION NUMBER: 08/482,711
; EARLIER FILING DATE: 1995-06-07
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 1475
; TYPE: PRT
; ORGANISM: Streptococcus mutans
US-09-210-361-2

```

```

Query Match          90.8%; Score 108; DB 3; Length 1475;
Best Local Similarity 85.7%; Pred. No. 1.3e-08;
Matches 18; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      1 ANNHVSIVEAMSDNDTPYLHD 21
Db      481 ANDHSLTEAMSDNDTPYLHD 501

```

```

RESULT 3
US-09-740-274-2
; Sequence 2, Application US/09740274
; Patent No. 6465203
; GENERAL INFORMATION:
; APPLICANT: Nichols, Scott E.
; TITLE OF INVENTION: Glucan-containing Compositions and Paper
; FILE REFERENCE: 0357CRD
; CURRENT APPLICATION NUMBER: US/09/740,274
; PRIOR FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: 09/210,361
; PRIOR FILING DATE: 1998-12-11
; PRIOR APPLICATION NUMBER: 09/007,999
; PRIOR FILING DATE: 1998-01-16
; PRIOR APPLICATION NUMBER: 08/478,704
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: 09/009,620
; PRIOR FILING DATE: 1998-01-20
; PRIOR APPLICATION NUMBER: 08/485,243
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: 09/008,172
; PRIOR FILING DATE: 1998-01-16
; PRIOR APPLICATION NUMBER: 08/482,711
; PRIOR FILING DATE: 1995-06-07
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 1475
; TYPE: PRT
; ORGANISM: Streptococcus mutans
US-09-740-274-2

```

```

Query Match          90.8%; Score 108; DB 4; Length 1475;
Best Local Similarity 85.7%; Pred. No. 1.3e-08;
Matches 18; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      1 ANNHVSIVEAMSDNDTPYLHD 21
Db      481 ANDHSLTEAMSDNDTPYLHD 501

```

```

RESULT 4
US-09-210-361-4
; Sequence 4, Application US/09210361
; Patent No. 6284479
; GENERAL INFORMATION:
; APPLICANT: Nichols, Scott E.
; TITLE OF INVENTION: Substitutes for Modified Starches and
; FILE REFERENCE: 0357CRD
; CURRENT APPLICATION NUMBER: US/09/210,361
; PRIOR FILING DATE: 1998-12-11
; EARLIER APPLICATION NUMBER: 09/007,999

```

```

; EARLIER FILING DATE: 1998-01-16
; EARLIER APPLICATION NUMBER: 08/478,704
; EARLIER FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 09/009,620
; EARLIER FILING DATE: 1998-01-20
; EARLIER APPLICATION NUMBER: 08/485,243
; EARLIER FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 09/008,172
; EARLIER FILING DATE: 1998-01-16
; EARLIER APPLICATION NUMBER: 08/482,711
; EARLIER FILING DATE: 1995-06-07
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 1375
; TYPE: PRT
; ORGANISM: streptococcus mutans
US-09-210-361-4

```

```

Query Match          83.2%; Score 99; DB 3; Length 1375;
Best Local Similarity 81.0%; Pred. No. 3.2e-07;
Matches 17; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

```

```

QY      1 ANNHVSIVEAMSDNDTPYLHD 21
Db      507 ANDHSLTEAMSYNDTPYLHD 527

```

```

RESULT 5
US-09-740-274-4
; Sequence 4, Application US/09740274
; Patent No. 6465203
; GENERAL INFORMATION:
; APPLICANT: Nichols, Scott E.
; TITLE OF INVENTION: Glucan-containing Compositions and Paper
; FILE REFERENCE: 0357CRD
; CURRENT APPLICATION NUMBER: US/09/740,274
; PRIOR FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: 09/210,361
; PRIOR FILING DATE: 1998-12-11
; PRIOR APPLICATION NUMBER: 09/007,999
; PRIOR FILING DATE: 1998-01-16
; PRIOR APPLICATION NUMBER: 08/478,704
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: 09/009,620
; PRIOR FILING DATE: 1998-01-20
; PRIOR APPLICATION NUMBER: 08/485,243
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: 09/008,172
; PRIOR FILING DATE: 1998-01-16
; PRIOR APPLICATION NUMBER: 08/482,711
; PRIOR FILING DATE: 1995-06-07
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 1375
; TYPE: PRT
; ORGANISM: streptococcus mutans
US-09-740-274-4

```

```

Query Match          83.2%; Score 99; DB 4; Length 1375;
Best Local Similarity 81.0%; Pred. No. 3.2e-07;
Matches 17; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

```

```

QY      1 ANNHVSIVEAMSDNDTPYLHD 21
Db      507 ANDHSLTEAMSYNDTPYLHD 527

```

```

RESULT 6
US-09-604-957-4
; Sequence 4, Application US/09604957
; Patent No. 6486314

```

```

GENERAL INFORMATION:
APPLICANT: VAN GEEL-SCHUTTEN, GERRITDINA HENDRIKA
APPLICANT: DIJKHUIZEN, LUBBERT
APPLICANT: RAHAOUI, HAKIM
APPLICANT: LEER, ROBERT-JAN
TITLE OF INVENTION: NOVEL GLUCOSYLTRANSFERASE AND NOVEL GLUCAN
FILE REFERENCE: BO 4338
CURRENT APPLICATION NUMBER: US/09/604,957
CURRENT FILING DATE: 2000-06-28
PRIOR APPLICATION NUMBER: 00201871.1
PRIOR FILING DATE: 2000-05-25
NUMBER OF SEQ ID NOS: 17
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 4
LENGTH: 545
TYPE: PRT
ORGANISM: Streptococcus mutans
US-09-604-957-4

Query Match
Best Local Similarity 62.2%; Score 74; DB 4; Length 545;
Best Local Similarity 66.7%; Pred. No. 0.00098;
Matches 14; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 1 ANNHVSTVEAWSNDPTPLHD 21
DB 75 AINHLSTLEAWSNDPQYNKD 95

RESULT 7
US-09-008-172-2
Sequence 2, Application US/09008172
Patent No. 6127602
GENERAL INFORMATION:
APPLICANT: Nichols, Scott E.
TITLE OF INVENTION: Substitutes for Modified Starch and
FILE REFERENCE: 0358D
CURRENT APPLICATION NUMBER: US/09/008,172
CURRENT FILING DATE: 1998-01-16
EARLIER APPLICATION NUMBER: 08/482,711
EARLIER FILING DATE: 1995-06-07
NUMBER OF SEQ ID NOS: 2
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 2
LENGTH: 1430
TYPE: PRT
ORGANISM: Streptococcus mutans
US-09-008-172-2

Query Match
Best Local Similarity 62.2%; Score 74; DB 3; Length 1430;
Best Local Similarity 66.7%; Pred. No. 0.003;
Matches 14; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 1 ANNHVSTVEAWSNDPTPLHD 21
DB 495 AINHLSTLEAWSNDPQYNKD 515

RESULT 8
US-09-210-361-6
Sequence 6, Application US/09210361
Patent No. 6284479
GENERAL INFORMATION:
APPLICANT: Nichols, Scott E.
TITLE OF INVENTION: Substitutes for Modified Starches and
FILE REFERENCE: 0357CR
CURRENT APPLICATION NUMBER: US/09/210,361
CURRENT FILING DATE: 1998-12-11
EARLIER APPLICATION NUMBER: 09/007,999
EARLIER FILING DATE: 1998-01-16
EARLIER APPLICATION NUMBER: 08/478,704
EARLIER FILING DATE: 1995-06-07

```

```

EARLIER APPLICATION NUMBER: 09/009,620
EARLIER FILING DATE: 1998-01-20
EARLIER APPLICATION NUMBER: 08/485,243
EARLIER FILING DATE: 1995-06-07
EARLIER APPLICATION NUMBER: 09/008,172
EARLIER FILING DATE: 1998-01-16
EARLIER APPLICATION NUMBER: 08/482,711
EARLIER FILING DATE: 1995-06-07
NUMBER OF SEQ ID NOS: 6
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 6
LENGTH: 1430
TYPE: PRT
ORGANISM: streptococcus mutans
US-09-210-361-6

Query Match
Best Local Similarity 62.2%; Score 74; DB 3; Length 1430;
Best Local Similarity 66.7%; Pred. No. 0.003;
Matches 14; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 1 ANNHVSTVEAWSNDPTPLHD 21
DB 495 AINHLSTLEAWSNDPQYNKD 515

RESULT 9
US-09-740-274-6
Sequence 6, Application US/09740274
Patent No. 6465203
GENERAL INFORMATION:
APPLICANT: Nichols, Scott E.
TITLE OF INVENTION: Glucan-containing Compositions and Paper
FILE REFERENCE: 0357CRD
CURRENT APPLICATION NUMBER: US/09/740,274
CURRENT FILING DATE: 2000-12-19
PRIOR APPLICATION NUMBER: 09/210,361
PRIOR FILING DATE: 1998-12-11
PRIOR APPLICATION NUMBER: 09/007,999
PRIOR FILING DATE: 1998-01-16
PRIOR APPLICATION NUMBER: 08/478,704
PRIOR FILING DATE: 1995-06-07
PRIOR APPLICATION NUMBER: 09/009,620
PRIOR FILING DATE: 1998-01-20
PRIOR APPLICATION NUMBER: 08/485,243
PRIOR FILING DATE: 1995-06-07
PRIOR APPLICATION NUMBER: 09/008,172
PRIOR FILING DATE: 1998-01-16
PRIOR APPLICATION NUMBER: 08/482,711
PRIOR FILING DATE: 1995-06-07
NUMBER OF SEQ ID NOS: 6
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 6
LENGTH: 1430
TYPE: PRT
ORGANISM: streptococcus mutans
US-09-740-274-6

Query Match
Best Local Similarity 62.2%; Score 74; DB 4; Length 1430;
Best Local Similarity 66.7%; Pred. No. 0.003;
Matches 14; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 1 ANNHVSTVEAWSNDPTPLHD 21
DB 495 AINHLSTLEAWSNDPQYNKD 515

RESULT 10
US-09-604-957-5
Sequence 5, Application US/09604957
Patent No. 6486314
GENERAL INFORMATION:
APPLICANT: VAN GEEL-SCHUTTEN, GERRITDINA HENDRIKA
APPLICANT: DIJKHUIZEN, LUBBERT

```

```

; APPLICANT: RAHAOUI, HAKIM
; APPLICANT: LEER, ROBERT-JAN
; TITLE OF INVENTION: NOVEL GLUCOSYLTRANSFERASE AND NOVEL GLUCAN
; FILE REFERENCE: BO 43388
; CURRENT APPLICATION NUMBER: US/09/604,957
; CURRENT FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: 00201871.1
; PRIOR FILING DATE: 2000-05-25
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 5
; LENGTH: 523
; TYPE: PRT
; ORGANISM: Leuconostoc mesenteroides
;
US-09-604-957-5

Query Match          54.6%; Score 65; DB 4; Length 523;
Best Local Similarity 57.1%; Pred. No. 0.025;
Matches 12; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY      1 ANNHVSIVAMSDNDTPYLHD 21
Db      75 ANQHLSILEDWMSHNDPLYVTD 95

RESULT 11
US-08-793-824-2
; Sequence 2, Application US/08793824
; Patent No. 5981838
; GENERAL INFORMATION:
; APPLICANT: Simpson, Christine Lynn
; APPLICANT: Giffard, Philip Morrison
; APPLICANT: Jacques, Nicholas Anthony
; TITLE OF INVENTION: Genetic Manipulation of Plants to
; TITLE OF INVENTION: Increase Stored Carbohydrates
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Griffith Hack & Co
; STREET: Level 8, 168 Walker Street
; CITY: No. 5981838th Sydney
; STATE: New South Wales
; COUNTRY: Australia
; ZIP: 2060
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/793,824
; FILING DATE:
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: AU PM7643
; FILING DATE: 24-AUG-1994
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 61 2 9957 5944
; TELEFAX: 61 2 957 6288
; TELEX: 26547
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1577 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: not relevant
; MOLECULE TYPE: Protein
; ORIGINAL SOURCE:
; ORGANISM: Streptococcus salivarius
;
US-08-793-824-2

Query Match          47.9%; Score 57; DB 2; Length 1577;
Best Local Similarity 61.1%; Pred. No. 1.6;
Matches 11; Conservative 2; Mismatches 5; Indels 0; Gaps 0;
```

```

QY      4 HVSIVAMSDNDTPYLHD 21
Db      591 HSLIEMSYNDHQYNKD 608

RESULT 12
US-09-604-957-6
; Sequence 6, Application US/09604957
; Patent No. 6486314
; GENERAL INFORMATION:
; APPLICANT: VAN GEEL-SCHUTTEN, GERRITJINA HENDRIKA
; APPLICANT: DIJKHUIZEN, LUBBERT
; APPLICANT: RAHAOUI, HAKIM
; APPLICANT: LEER, ROBERT-JAN
; TITLE OF INVENTION: NOVEL GLUCOSYLTRANSFERASE AND NOVEL GLUCAN
; FILE REFERENCE: BO 43388
; CURRENT APPLICATION NUMBER: US/09/604,957
; CURRENT FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: 00201871.1
; PRIOR FILING DATE: 2000-05-25
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 6
; LENGTH: 584
; TYPE: PRT
; ORGANISM: Leuconostoc mesenteroides
;
US-09-604-957-6

Query Match          46.2%; Score 55; DB 4; Length 584;
Best Local Similarity 45.0%; Pred. No. 1.1;
Matches 9; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

QY      1 ANNHVSIVAMSDNDTPYLH 20
Db      75 ANKHLSILEDWNGKDPQYVN 94

RESULT 13
US-09-499-203-2
; Sequence 2, Application US/09499203
; Patent No. 6570065
; GENERAL INFORMATION:
; APPLICANT: KOSSMANN, Jens
; APPLICANT: WELSH, Thomas
; APPLICANT: QUANZ, Martin
; APPLICANT: KNUTH, Karola
; TITLE OF INVENTION: Nucleic Acid Molecules Encoding Alternansucrase
; FILE REFERENCE: 147-196P
; CURRENT APPLICATION NUMBER: US/09/499,203
; CURRENT FILING DATE: 2000-02-08
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 2
; LENGTH: 2057
; TYPE: PRT
; ORGANISM: Leuconostoc mesenteroides
;
US-09-499-203-2

Query Match          46.2%; Score 55; DB 4; Length 2057;
Best Local Similarity 45.0%; Pred. No. 4.5;
Matches 9; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

QY      1 ANNHVSIVAMSDNDTPYLH 20
Db      665 ANKHLSILEDWNGKDPQYVN 684

RESULT 14
US-09-604-957-7
; Sequence 7, Application US/09604957
; Patent No. 6486314
; GENERAL INFORMATION:
```



```

; APPLICANT: VAN GEEL-SCHUTTEN, GERRITDINA HENDRIKA
; APPLICANT: DIJKHUIZEN, LUBBERT
; APPLICANT: RAHAOUI, HAKIM
; APPLICANT: LEER, ROBERT-JAN
; TITLE OF INVENTION: NOVEL GLUCOSYLTRANSFERASE AND NOVEL GLUCAN
; FILE REFERENCE: BO 43388
; CURRENT APPLICATION NUMBER: US/09/604,957
; CURRENT FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: 00201871.1
; PRIOR FILING DATE: 2000-05-25
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 7
; LENGTH: 535
; TYPE: PRT
; ORGANISM: Lactobacillus reuteri
US-09-604-957-7

```

```

Query Match          40.3%; Score 48; DB 4; Length 535;
Best Local Similarity 35.0%; Pred. No. 12;
Matches 7; Conservative 6; Mismatches 7; Indels 0; Gaps 0;

```

```

Qy      1 ANHWSTVEAMSDNDTPYLH 20
Db      74 SNKHINILEDWNNHADPEYFN 93

```

```

RESULT 15
US-09-604-957-3
; Sequence 3, Application US/09604957
; Patent No. 6486314
; GENERAL INFORMATION:
; APPLICANT: VAN GEEL-SCHUTTEN, GERRITDINA HENDRIKA
; APPLICANT: DIJKHUIZEN, LUBBERT
; APPLICANT: RAHAOUI, HAKIM
; APPLICANT: LEER, ROBERT-JAN
; TITLE OF INVENTION: NOVEL GLUCOSYLTRANSFERASE AND NOVEL GLUCAN
; FILE REFERENCE: BO 43388
; CURRENT APPLICATION NUMBER: US/09/604,957
; CURRENT FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: 00201871.1
; PRIOR FILING DATE: 2000-05-25
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 1278
; TYPE: PRT
; ORGANISM: Lactobacillus reuteri
US-09-604-957-3

```

```

Query Match          40.3%; Score 48; DB 4; Length 1278;
Best Local Similarity 35.0%; Pred. No. 34;
Matches 7; Conservative 6; Mismatches 7; Indels 0; Gaps 0;

```

```

Qy      1 ANHWSTVEAMSDNDTPYLH 20
Db      550 SNKHINILEDWNNHADPEYFN 569

```

Search completed: November 13, 2003, 09:48:00
Job time : 14.6351 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using SW model

Run on: November 13, 2003, 09:45:40 ; Search time 27.4692 seconds
(without alignments)
139.565 Million cell updates/sec

Title: US-09-290-049a-12
Perfect score: 119
Sequence: 1 ANNHVSIVEAWSNDTPYLHD 21

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 666188 seqs, 182559486 residues

Total number of hits satisfying chosen parameters: 666188

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database:

1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/2/pubpaa/PTCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/2/pubpaa/PTCT_PUBCOMB.pep:*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep:*
10: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep:*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep:*
13: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep:*
14: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep:*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
17: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match length	ID	Description
1	108	90.8	1475	US-09-740-274-2
2	99	83.2	1375	US-09-740-274-4
3	74	62.2	545	US-09-995-749A-10
4	74	62.2	1430	US-09-740-274-6
5	54.6	522	10	US-09-995-749A-11
6	46.2	584	10	US-09-995-749A-12
7	48	40.3	535	US-09-995-749A-13
8	48	40.3	1781	US-09-995-749A-2
9	47	39.5	749	US-10-211-962-96
10	46.5	39.1	1604	US-09-888-615-73
11	46	38.7	502	US-10-009-823A-1
12	46	38.7	1494	US-10-195-144-17
13	45	37.8	199	US-09-738-626-4571
14	45	37.8	665	US-10-156-761-14923
15	45	37.8	972	US-09-924-154-16

15	44	37.0	15	12	US-10-295-693-90	Sequence 90, Appl
17	44	37.0	15	15	US-10-067-649-90	Sequence 90, Appl
18	43.5	36.6	202	10	US-09-280-197-11	Sequence 11, Appl
19	43.5	36.6	202	12	US-10-448-139-11	Sequence 11, Appl
20	43.5	36.6	491	9	US-09-815-242-10190	Sequence 10190, A
21	43.5	36.6	491	9	US-09-815-242-13756	Sequence 13756, A
22	43.5	36.6	1092	10	US-09-423-126-5	Sequence 5, Appl
23	43	36.1	27	15	US-10-092-908-26	Sequence 26, Appl
24	43	36.1	67	15	US-10-050-704-234	Sequence 324, Appl
25	43	36.1	87	15	US-10-050-704-149	Sequence 149, Appl
26	43	36.1	332	10	US-09-454-279-20	Sequence 20, Appl
27	43	36.1	437	11	US-09-374-046A-66	Sequence 66, Appl
28	43	36.1	539	13	US-10-230-026-34	Sequence 34, Appl
29	43	36.1	603	15	US-10-054-678-35	Sequence 2, Appl
30	43	36.1	603	15	US-10-092-908-35	Sequence 35, Appl
31	43	36.1	715	15	US-10-156-761-7562	Sequence 7562, Ap
32	42	35.3	349	12	US-10-220-381-22	Sequence 22, Appl
33	42	35.3	432	9	US-09-815-242-5096	Sequence 5096, Ap
34	42	35.3	455	9	US-09-742-954-6	Sequence 6, Appl
35	42	35.3	7349	12	US-10-314-657-46	Sequence 28981, A
36	41.5	34.9	49	12	US-10-029-386-28981	Sequence 146, App
37	41	34.5	120	14	US-10-001-835-146	Sequence 146, App
38	41	34.5	302	9	US-09-815-242-14004	Sequence 14004, A
39	41	34.5	358	11	US-09-095-478-7	Sequence 9, Appl
40	41	34.5	381	11	US-09-095-478-7	Sequence 7, Appl
41	41	34.5	438	15	US-10-156-761-14114	Sequence 14114, A
42	41	34.5	729	15	US-10-156-761-11729	Sequence 11729, A
43	41	34.5	1267	15	US-10-059-585-56	Sequence 56, Appl
44	41	34.5	2466	12	US-10-177-980-12	Sequence 12, Appl
45	41	34.5	2485	9	US-09-802-669-46	Sequence 46, Appl

ALIGNMENTS

RESULT 1
US-09-740-274-2
Sequence 2, Application US/09740274
Patent No. US20020031826A1

GENERAL INFORMATION:
APPLICANT: Nichols, Scott E.
TITLE OF INVENTION: Glucan-containing Compositions and Paper
FILE REFERENCE: 0357CRD

CURRENT FILING DATE: 2000-12-19 274

PRIOR FILING DATE: 1998-12-11 361

PRIOR FILING DATE: 1998-01-16 999

PRIOR FILING DATE: 1998-01-16 704

PRIOR FILING DATE: 1995-06-07 620

PRIOR FILING DATE: 1998-01-20 243

PRIOR FILING DATE: 1995-06-07 172

PRIOR FILING DATE: 1998-01-16 711

PRIOR FILING DATE: 1995-06-07 6

NUMBER OF SEQ ID NOS: 6

SOFTWARE: FastSeq for Windows Version 3.0

SEQ ID NO 2

LENGTH: 1475

TYPE: PRT

ORGANISM: Streptococcus mutans

US-09-740-274-2

Query Match 90.8%; Score 108; DB 9; Length 1475;
Best Local Similarity 85.7%; Pred. No. 7.3e-08;
Matches 18; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
1 ANNHVSIVEAWSNDTPYLHD 21
||:|||||

DB 481 ANDHLSILEAWSNDNTPYLHD 501

RESULT 2

US-09-740-274-4
; Sequence 4, Application US/09740274
; Patent No. US20020031826A1
; GENERAL INFORMATION:
; APPLICANT: NICHOLS, Scott E.
; TITLE OF INVENTION: Glucan-containing Compositions and Paper
; FILE REFERENCE: 0357CRD
; CURRENT APPLICATION NUMBER: US/09/740,274
; CURRENT FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: 09/210,361
; PRIOR FILING DATE: 1998-12-11
; PRIOR APPLICATION NUMBER: 09/007,999
; PRIOR FILING DATE: 1998-01-16
; PRIOR APPLICATION NUMBER: 08/478,704
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: 09/009,620
; PRIOR FILING DATE: 1998-01-20
; PRIOR APPLICATION NUMBER: 08/485,243
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: 09/008,172
; PRIOR FILING DATE: 1998-01-16
; PRIOR APPLICATION NUMBER: 08/482,711
; PRIOR FILING DATE: 1995-06-07
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 1375
; TYPE: PRT
; ORGANISM: streptococcus mutans
US-09-740-274-4

Query Match 83.2%; Score 99; DB 9; Length 1375;
Best Local Similarity 81.0%; Pred. No. 1,6e-06;
Matches 17; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 ANNHVSIWEAWSNDNTPYLHD 21

DB 507 ANDHLSILEAWSNDNTPYLHD 527

RESULT 3
US-09-995-749A-10
; Sequence 10, Application US/09995749A
; Patent No. US20020155568A1
; GENERAL INFORMATION:
; APPLICANT: VAN GEEL-SCHUTTEN, GERRITDINA HENDRIKA
; APPLICANT: DIJKHUIZEN, LUBBERT
; APPLICANT: RAHOUTI, HAKIM
; APPLICANT: LEER, ROBERT-JAN
; TITLE OF INVENTION: NOVEL GLUCOSYLTRANSFERASES
; FILE REFERENCE: B043388-CIP
; CURRENT APPLICATION NUMBER: US/09/995,749A
; CURRENT FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: 09/604,957
; PRIOR FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: EPO 00201871.1
; PRIOR FILING DATE: 2000-05-25
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
; LENGTH: 545
; TYPE: PRT
; ORGANISM: Streptococcus mutans
US-09-995-749A-10

Query Match 62.2%; Score 74; DB 10; Length 545;
Best Local Similarity 66.7%; Pred. No. 0.0042;
Matches 14; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 1 ANNHVSIWEAWSNDNTPYLHD 21
DB 75 ANHLSILEAWSNDNTPQYNKD 95

RESULT 4

US-09-740-274-6
; Sequence 6, Application US/09740274
; Patent No. US20020031826A1
; GENERAL INFORMATION:
; APPLICANT: NICHOLS, Scott E.
; TITLE OF INVENTION: Glucan-containing Compositions and Paper
; FILE REFERENCE: 0357CRD
; CURRENT APPLICATION NUMBER: US/09/740,274
; CURRENT FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: 09/210,361
; PRIOR FILING DATE: 1998-12-11
; PRIOR APPLICATION NUMBER: 09/007,999
; PRIOR FILING DATE: 1998-01-16
; PRIOR APPLICATION NUMBER: 08/478,704
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: 09/009,620
; PRIOR FILING DATE: 1998-01-20
; PRIOR APPLICATION NUMBER: 08/485,243
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: 09/008,172
; PRIOR FILING DATE: 1998-01-16
; PRIOR APPLICATION NUMBER: 08/482,711
; PRIOR FILING DATE: 1995-06-07
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 6
; LENGTH: 1430
; TYPE: PRT
; ORGANISM: streptococcus mutans
US-09-740-274-6

Query Match 62.2%; Score 74; DB 9; Length 1430;
Best Local Similarity 66.7%; Pred. No. 0.012;
Matches 14; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 1 ANNHVSIWEAWSNDNTPYLHD 21

DB 495 ANHLSILEAWSNDNTPQYNKD 515

RESULT 5
US-09-995-749A-11
; Sequence 11, Application US/09995749A
; Patent No. US20020155568A1
; GENERAL INFORMATION:
; APPLICANT: VAN GEEL-SCHUTTEN, GERRITDINA HENDRIKA
; APPLICANT: DIJKHUIZEN, LUBBERT
; APPLICANT: RAHOUTI, HAKIM
; APPLICANT: LEER, ROBERT-JAN
; TITLE OF INVENTION: NOVEL GLUCOSYLTRANSFERASES
; FILE REFERENCE: B043388-CIP
; CURRENT APPLICATION NUMBER: US/09/995,749A
; CURRENT FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: 09/604,957
; PRIOR FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: EPO 00201871.1
; PRIOR FILING DATE: 2000-05-25
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 11
; LENGTH: 522
; TYPE: PRT
; ORGANISM: Leuconostoc mesenteroides
US-09-995-749A-11

Query Match 54.6%; Score 65; DB 10; Length 522;
Best Local Similarity 57.1%; Pred. No. 0.097;

Matches 12; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 1 ANNHVSIVEAMSDNDTPYLH 21
DB 75 ANGHLSLEDMNDPQYVN 95

RESULT 6
US-09-995-749A-12
; Sequence 12, Application US/09995749A
; Patent No. US20020155568A1
; GENERAL INFORMATION:
; APPLICANT: VAN GEEL-SCHUTTEN, GERRITDINA HENDRIKA
; APPLICANT: DIJKHUIZEN, LOUBBERT
; APPLICANT: RAHAOUI, HAKIM
; APPLICANT: LEER, ROBERT-JAN
; TITLE OF INVENTION: NOVEL GLUCOSYLTRANSFERASES
; FILE REFERENCE: B043388-CIP
; CURRENT APPLICATION NUMBER: US/09/995,749A
; PRIOR FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: 09/604,957
; PRIOR FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: EPO 00201871.1
; PRIOR FILING DATE: 2000-05-25
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 12
; LENGTH: 584
; TYPE: PRT
; ORGANISM: Leuconobloc mesenteroides
US-09-995-749A-12

Query Match 46.2%; Score 55; DB 10; Length 584;
Best Local Similarity 45.0%; Pred. No. 3.8;
Matches 9; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

QY 1 ANNHVSIVEAMSDNDTPYLH 20
DB 75 ANKHLSLEDMNDPQYVN 94

RESULT 7
US-09-995-749A-13
; Sequence 13, Application US/09995749A
; Patent No. US20020155568A1
; GENERAL INFORMATION:
; APPLICANT: VAN GEEL-SCHUTTEN, GERRITDINA HENDRIKA
; APPLICANT: DIJKHUIZEN, LOUBBERT
; APPLICANT: RAHAOUI, HAKIM
; APPLICANT: LEER, ROBERT-JAN
; TITLE OF INVENTION: NOVEL GLUCOSYLTRANSFERASES
; FILE REFERENCE: B043388-CIP
; CURRENT APPLICATION NUMBER: US/09/995,749A
; CURRENT FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: 09/604,957
; PRIOR FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: EPO 00201871.1
; PRIOR FILING DATE: 2000-05-25
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 13
; LENGTH: 535
; TYPE: PRT
; ORGANISM: Lactobacillus reuteri
US-09-995-749A-13

Query Match 40.3%; Score 48; DB 10; Length 535;
Best Local Similarity 35.0%; Pred. No. 4.1;
Matches 7; Conservative 6; Mismatches 7; Indels 0; Gaps 0;

QY 1 ANNHVSIVEAMSDNDTPYLH 20
DB 74 SNKHLSLEDMNDPQYVN 93

RESULT 8
US-09-995-749A-2
; Sequence 2, Application US/09995749A
; Patent No. US20020155568A1
; GENERAL INFORMATION:
; APPLICANT: VAN GEEL-SCHUTTEN, GERRITDINA HENDRIKA
; APPLICANT: DIJKHUIZEN, LOUBBERT
; APPLICANT: RAHAOUI, HAKIM
; APPLICANT: LEER, ROBERT-JAN
; TITLE OF INVENTION: NOVEL GLUCOSYLTRANSFERASES
; FILE REFERENCE: B043388-CIP
; CURRENT APPLICATION NUMBER: US/09/995,749A
; CURRENT FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: 09/604,957
; PRIOR FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: EPO 00201871.1
; PRIOR FILING DATE: 2000-05-25
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 1781
; TYPE: PRT
; ORGANISM: Lactobacillus reuteri
US-09-995-749A-2

Query Match 40.3%; Score 48; DB 10; Length 1781;
Best Local Similarity 35.0%; Pred. No. 1.6e+02;
Matches 7; Conservative 6; Mismatches 7; Indels 0; Gaps 0;

QY 1 ANNHVSIVEAMSDNDTPYLH 20
DB 1053 SNKHLSLEDMNDPQYVN 1072

RESULT 9
US-10-211-962-96
; Sequence 96, Application US/10211962
; Publication No. US20030082640A1
; GENERAL INFORMATION:
; APPLICANT: Herz, Joachim
; APPLICANT: Gotthardt, Michael
; TITLE OF INVENTION: LDL Receptor Signaling Pathways
; FILE REFERENCE: UTSW0708
; CURRENT APPLICATION NUMBER: US/10/211,962
; CURRENT FILING DATE: 2002-08-01
; PRIOR APPLICATION NUMBER: US/09/562,737
; PRIOR FILING DATE: 2000-05-01
; NUMBER OF SEQ ID NOS: 132
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 96
; LENGTH: 749
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: Sequence
US-10-211-962-96

Query Match 39.5%; Score 47; DB 15; Length 749;
Best Local Similarity 53.3%; Pred. No. 86;
Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 6 SIVEAMSDNDTPYLH 20
DB 136 SAVEEMTDSATPHPH 150

RESULT 10
US-09-888-615-73
; Sequence 73, Application US/09888615
; Patent No. US20020064856A1

```

; GENERAL INFORMATION:
; APPLICANT: PLOWMAN, GREGORY
; APPLICANT: MWAYTE, DAVID
; APPLICANT: CAENEPEEL, SEAN
; APPLICANT: CHAYDOZAK, GLEN
; APPLICANT: MANNING, GERARD
; APPLICANT: SUDARSANAM, SUCHA
; TITLE OF INVENTION: NOVEL PROTEASES
; FILE REFERENCE: 038602/1214
; CURRENT APPLICATION NUMBER: US/09/888,615
; PRIOR FILING DATE: 2001-06-26
; PRIOR APPLICATION NUMBER: 60/214,047
; PRIOR FILING DATE: 2000-06-26
; NUMBER OF SEQ ID NOS: 150
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 73
; LENGTH: 1604
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-888-615-73

```

```

Query Match          39.1%; Score 46.5; DB 9; Length 1604;
Best Local Similarity 47.4%; Pred. No. 2.4e+02;
Matches 9; Conservative 3; Mismatches 4; Indels 3; Gaps 1;
QY 5 VSIVEAMSDN--DTPYLH 20
DB 294 VALLEWKNRRTDIDPELH 312

```

```

RESULT 11
US-10-009-823A-1
; Sequence 1, Application US/10009823A
; Publication No. US20030157120A1
; GENERAL INFORMATION:
; APPLICANT: Panaccio, Michael
; APPLICANT: Rosey, Everett Lee
; APPLICANT: Sinistaj, Merti
; APPLICANT: Hasse, Detlef
; APPLICANT: Parsons, Jim
; APPLICANT: Ankenbauer, Robert G.
; TITLE OF INVENTION: LAMSONIA DERIVED GENE AND RELATED FLGE
; TITLE OF INVENTION: POLYPEPTIDES, PEPTIDES AND PROTEINS AND THEIR USES
; FILE REFERENCE: DAV1150.001APC
; CURRENT APPLICATION NUMBER: US/10/009,823A
; PRIOR FILING DATE: 2002-08-13
; PRIOR APPLICATION NUMBER: PCT/AU00/00437
; PRIOR FILING DATE: 2000-05-11
; PRIOR APPLICATION NUMBER: US 60/133,973
; PRIOR FILING DATE: 1999-05-13
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 502
; TYPE: PRT
; ORGANISM: Lawsonia intracellularis
US-10-009-823A-1

```

```

Query Match          38.7%; Score 46; DB 12; Length 502;
Best Local Similarity 41.2%; Pred. No. 79;
Matches 7; Conservative 5; Mismatches 5; Indels 0; Gaps 0;
QY 1 ANNHVSIVEAMSDNDTP 17
DB 185 ANPYFALLESWKNGTTP 201

```

```

RESULT 12
US-10-195-144-17
; Sequence 17, Application US/10195144
; Publication No. US2003012646A1
; GENERAL INFORMATION:
; APPLICANT: BROWN, GREGORY G.

```

```

; APPLICANT: FORMANOVA, NATASA
; APPLICANT: DENDY, CHARLES
; APPLICANT: LANDRY, BENOTT S.
; APPLICANT: CHEUNG, WING
; APPLICANT: JIN, HUA
; TITLE OF INVENTION: NUCLEAR FERTILITY RESTORER GENES AND METHODS OF USE IN
; FILE REFERENCE: 16313-0136
; CURRENT APPLICATION NUMBER: US/10/195,144
; PRIOR FILING DATE: 2002-10-01
; PRIOR APPLICATION NUMBER: 60/305,026
; PRIOR FILING DATE: 2001-07-12
; PRIOR APPLICATION NUMBER: 60/305,363
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: 60/308,736
; PRIOR FILING DATE: 2001-07-30
; NUMBER OF SEQ ID NOS: 128
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 17
; LENGTH: 1494
; TYPE: PRT
; ORGANISM: Raphanus sativum
US-10-195-144-17

```

```

Query Match          38.7%; Score 46; DB 15; Length 1494;
Best Local Similarity 41.2%; Pred. No. 2.7e+02;
Matches 7; Conservative 3; Mismatches 7; Indels 0; Gaps 0;
QY 3 NHVSIVEAMSDNDTPYL 19
DB 1130 NTISLAFTWLNQSEPL 1146

```

```

RESULT 13
US-09-738-626-4571
; Sequence 4571, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIRO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAOKO
; APPLICANT: SENOH, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738,626
; PRIOR FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: PatentIn Ver. 3.0
; SEQ ID NO 4571
; LENGTH: 199
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-738-626-4571

```

```

Query Match          37.8%; Score 45; DB 10; Length 199;
Best Local Similarity 58.3%; Pred. No. 40;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
QY 7 IVEAMSDNDTPY 18
DB 177 LVEWTDQDTRY 188

```

RESULT 14

US-10-156-761-14923
 ; Sequence 14923, Application US/10156761
 ; Publication No. US20030119018A1
 ; GENERAL INFORMATION:
 ; APPLICANT: OMURA, SATOSHI
 ; APPLICANT: IKEDA, HARUO
 ; APPLICANT: ISHIKAWA, JUN
 ; APPLICANT: HORIKAWA, HIROSHI
 ; APPLICANT: SHIBA, TADAYOSHI
 ; APPLICANT: SAKAKI, YOSHIYUKI
 ; APPLICANT: HATTORI, MASAHIRO
 ; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
 ; FILE REFERENCE: 249-262
 ; CURRENT APPLICATION NUMBER: US/10/156,761
 ; CURRENT FILING DATE: 2002-05-29
 ; PRIOR APPLICATION NUMBER: JP 2001-204089
 ; PRIOR FILING DATE: 2001-05-30
 ; PRIOR APPLICATION NUMBER: JP 2001-272697
 ; PRIOR FILING DATE: 2001-08-02
 ; NUMBER OF SEQ ID NOS: 15109
 ; SEQ ID NO 14923
 ; LENGTH: 665
 ; TYPE: PRT
 ; ORGANISM: Streptomyces avermitilis
 US-10-156-761-14923

Query Match

Best Local Similarity 37.8%; Score 45; DB 15; Length 665;
 Pred. No. 1.5e+02;
 Matches 9; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

OY 3 NHVSIVEAWSNDPTPYLH 20
 | : ||||| : |

DB 88 NMTYVEAWSDPYTTWRH 105

RESULT 15

US-09-924-154-16
 ; Sequence 16, Application US/09924154
 ; Patent No. US20020127241A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Narum, David L.
 ; TITLE OF INVENTION: Anti-Plasmodium Compositions and Methods of Use
 ; FILE REFERENCE: 05213-0465 43170-262105
 ; CURRENT APPLICATION NUMBER: US/09/924,154
 ; CURRENT FILING DATE: 2001-08-07
 ; PRIOR APPLICATION NUMBER: US 60/223,525
 ; PRIOR FILING DATE: 2000-08-07
 ; NUMBER OF SEQ ID NOS: 17
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 16
 ; LENGTH: 972
 ; TYPE: PRT
 ; ORGANISM: Mammalian
 US-09-924-154-16

Query Match

Best Local Similarity 37.8%; Score 45; DB 10; Length 972;
 Pred. No. 2.3e+02;
 Matches 8; Conservative 1; Mismatches 9; Indels 0; Gaps 0;

OY 3 NHVSIVEAWSNDPTPYLH 20
 | : ||||| : |

DB 304 NHTGVCLMEDDNDNGYLH 321

Search completed: November 13, 2003, 10:29:03
 Job time : 28.4692 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 13, 2003, 09:11:40 ; Search time 38.5166 Seconds
(without alignments)
86.541 Million cell updates/sec

Title: US-09-290-049A-13
Perfect score: 113
Sequence: 1 AIDHLSILEAMSGNDNDYKQ 21

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

```

1: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
8: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
11: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*
24: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	75	66.4	1430 23 AAU98029	S. mutans glucosyl
2	75	66.4	1430 23 AAU98041	S. mutans glucosyl
3	75	66.4	1430 23 AAU98042	S. mutans glucosyl
4	75	66.4	1430 23 AAU98043	S. mutans glucosyl
5	75	66.4	1430 23 AAU98044	S. mutans glucosyl
6	75	66.4	1430 23 AAU98045	S. mutans glucosyl
7	73	64.6	1017 23 AAU79285	Streptococcus muta
8	73	64.6	1475 23 AAU98027	S. mutans glucosyl
9	73	64.6	1475 23 AAU98030	S. mutans glucosyl

10	73	64.6	1475 23 AAU98031	S. mutans glucosyl
11	73	64.6	1475 23 AAU98032	S. mutans glucosyl
12	73	64.6	1475 23 AAU98033	S. mutans glucosyl
13	73	64.6	1475 23 AAU98034	S. mutans glucosyl
14	73	64.6	1475 23 AAU98035	S. mutans glucosyl
15	73	64.6	1475 23 AAU98036	S. mutans glucosyl
16	73	64.6	1475 23 AAU98037	S. mutans glucosyl
17	73	64.6	1475 23 AAU98038	S. mutans glucosyl
18	73	64.6	1475 23 AAU98039	S. mutans glucosyl
19	73	64.6	1475 23 AAU98040	S. mutans glucosyl
20	73	64.6	1475 23 AAU98041	S. mutans glucosyl
21	73	64.6	1577 17 AAU91047	Alpha-D-glucosyltr
22	71	62.8	1375 23 AAU98028	S. mutans glucosyl
23	71	62.8	1375 23 AAU98028	Streptococcus muta
24	66	58.4	12 23 ABB98643	Dextrane-saccharas
25	65	57.5	2057 21 AAB10667	L. mesenteroides a
26	64	56.6	2835 23 ABB98574	Dextran saccharase
27	62	54.9	1592 14 AAR32925	Glucosyltransferase
28	59	52.2	12 23 ABB98641	Dextrane-saccharas
29	58	51.3	1527 23 AAU80055	Leuconostoc mesent
30	53	46.9	12 23 ABB98642	Dextrane-saccharas
31	52	46.0	12 23 ABB98644	Dextrane-saccharas
32	52	46.0	12 23 ABB98647	Dextrane-saccharas
33	51	45.1	12 23 ABB98647	Dextrane-saccharas
34	50	44.4	121 22 AAU60582	Protonibacterium
35	49	43.4	1429 23 ABB47797	Yeast DNA topoisom
36	48.5	42.9	218 21 AAB43277	Human ORFX ORF3041
37	48.5	42.9	289 22 AAU25871	Human protein sequ
38	48.5	42.9	319 22 AAU2156	Novel human secret
39	48.5	42.9	427 20 AAU25761	Human secreted pro
40	48.5	42.9	427 21 AAB32412	Human secreted pro
41	48.5	42.9	436 21 AAB32411	Human secreted pro
42	48.5	42.9	576 22 AAB32384	Human secreted pro
43	48.5	42.9	576 22 AAB34297	Human protein sequ
44	48.5	42.9	576 22 ABB4569	Human protein SEO
45	48	42.5	12 23 ABB98578	Dextrane saccharas

ALIGNMENTS

RESULT 1	
AAU98029	AAU98029 standard; Protein; 1430 AA.
XX	
AC	AAU98029;
XX	
DT	27-AUG-2002 (first entry)
XX	
DE	S. mutans glucosyltransferase GTFP.
XX	
KW	Glucosyltransferase; GTFP; transgenic plant; paper sizing;
KW	coating composition; glucan; starch; latex; thermoplastic molecule;
KW	amylolast; vacuole; paper manufacture.
XX	
OS	Streptococcus mutans.
XX	
PN	US2002031826-A1.
XX	
PD	14-MAR-2002.
XX	
PF	19-DEC-2000; 2000US-0740274.
XX	
PR	11-DEC-1998; 98US-0210361.
PR	07-JUN-1995; 95US-0478704.
PR	07-JUN-1995; 95US-0482711.
PR	07-JUN-1995; 95US-0485243.
PR	16-JAN-1998; 98US-0007999.
PR	16-JAN-1998; 98US-0008172.
PR	20-JAN-1998; 98US-0009620.
XX	
PA	(NICH/) NICHOLS S E.
XX	

XX Nichols SE;
 XI WPI; 2002-414332/44.
 DR N-PSDB; ABR52940.
 XX
 PT Glucosyltransferase B or D protein useful for producing a glucan useful
 PR as substitutes for and additions to modified starch and latexes in
 PT paper manufacture, comprises mutations in specific positions -
 XX
 PS Disclosure; Page 38-42; 44pp; English.
 XX
 XX The invention an isolated protein comprising a glucosyltransferase
 CC (GTF) B polypeptide having changes at position from 1448V, D457N,
 CC D567T, K1014T, D457N/D567T, D457N/D571K, D567T/D571K,
 CC D567T/D571K/K1014T, 1448V/D457N/D567T/D571K/K779Q/K1014T,
 CC K163A/K170A/K171A, and K779Q or a GTF D polypeptide having
 CC changes at positions from T589D, T583E, N471D, N471D/T588D, and
 CC N471D/T589E. Also included are a glucan produced by the GTF mutant,
 CC an isolated polynucleotide which encodes P1 or P2, or its complementary
 CC polynucleotide, a ribonucleic acid sequence encoding the GTF mutant,
 CC an expression cassette comprising the polynucleotide operably linked to a
 CC promoter, a vector comprising the expression cassette, host cell
 CC introduced with the vector, a transgenic plant comprising the
 CC vector, a seed or tuber from the transgenic plant, a paper sizing and/or
 CC coating composition comprising a glucan produced in a plant transformed
 CC with a gene encoding the mutant GTF, wild type or, starch, a latex,
 CC thermoplastic molecule or their combinations or glucan and starch where
 CC the glucan is produced in the amyloplast and/or vacuole or a maize line
 CC deficient in starch biosynthesis, transformed with a gene encoding a
 CC glucosyltransferase B or D enzyme, wild-type or mutant and a paper
 CC comprising the glucan (paper sizing/coating agent). The vector is useful
 CC for producing a glucan in a plant. The method comprises transforming a
 CC plant cell with the vector, growing the plant cell under plant growing
 CC conditions to produce a regenerated plant and inducing expression of the
 CC polynucleotide for a time sufficient to produce the glucan in the
 CC regenerated plant, where the vector contains a transit sequence from
 CC ribulose biphosphate carboxylase small subunit, waxy, brittle-1 and
 CC chlorophyll AB binding protein to produce a transgenic plant, and glucan
 CC is produced in the amyloplast of potato or the vacuole of sugar beet.
 CC Glucans are useful as substitutes for and additions to modified starch
 CC and latexes in paper manufacture. Unlike prior art techniques, which
 CC require input materials that produce chemical effluents, paper
 CC manufacture utilizing the glucan produced by GTF, which utilizes
 CC biologically produced input materials, is more cost-effective and
 CC environmentally friendly. Moreover, glucans also exhibit thermoplastic
 CC properties and impart gloss to the paper during coating step.
 CC The present sequence represents GTFD.
 XX
 SO Sequence 1430 AA;
 QY 1 AIDHSLTLEAWSGNDNDYVK 20
 DQ 495 AINHLSTLEAWSGNDNDPYNK 514
 XX
 XX RESULT 2
 XX AAU98041
 XX ID AAU98041 standard; Protein; 1430 AA.
 XX AC AAU98041;
 XX DT 27-AUG-2002 (first entry)
 XX DE S. mutans glucosyltransferase GTFD mutant T589D.
 XX KW Glucosyltransferase; GTFD; transgenic plant; paper sizing;
 XX coating composition; glucan; starch; latex; thermoplastic molecule;
 XX amyloplast; vacuole; paper manufacture; mutant; muten.

OS	Streptococcus mutans.
XX	Synthetic.
FT	Key
XX	Location/Qualifiers
FT	Misc-difference 589
XX	/note= "Wild-type Thr substituted by Asp"
PN	US2002031826-A1.
XX	
PD	14-MAR-2002.
XX	
PF	19-DEC-2000; 2000US-0740274.
XX	
PR	11-DEC-1998; 98US-0210361.
XX	
PR	07-JUN-1995; 95US-0478704.
XX	
PR	07-JUN-1995; 95US-0482711.
XX	
PR	07-JUN-1995; 95US-0485243.
XX	
PR	16-JAN-1998; 98US-0007999.
XX	
PR	16-JAN-1998; 98US-0008172.
XX	
PR	20-JAN-1998; 98US-0009620.
XX	
PA	(NICH/) NICHOLS S E.
XX	
PI	Nichols SE.
XX	
DR	WPI; 2002-414332/44.
XX	
PT	Glucosyltransferase B or D protein useful for producing a glucan useful
XX	as substitutes for and additions to modified starch and latexes in
PT	paper manufacture, comprises mutations in specific positions -
XX	
PS	Claim 36; Page -: 44pp; English.
XX	
CC	The invention an isolated protein comprising a glucosyltransferase
CC	(GTF) B polypeptide having changes at position from 1448V, D457N,
CC	D567T, K1014T, D457N/D567T, D457N/D571K, D567T/D571K,
CC	D567T/D571K/K1014Y, 1448V/D457N/D567T/D571K/K779Q/K1014T,
CC	Y168A/Y170A/Y171A, and K779Q or a GTF D polypeptide having
CC	changes at positions from T589D, T589E, N471D, N471D/T589D, and
CC	N471D/T589E. Also included are a glucan produced by the GTF mutant,
CC	an isolated polynucleotide which encodes P1 or P2, or its complementary
CC	polynucleotide, a ribonucleic acid sequence encoding the GTF mutant,
CC	an expression cassette comprising the polynucleotide operably linked to a
CC	promoter, a vector comprising the expression cassette, host cell
CC	introduced with the vector, a transgenic plant comprising the
CC	vector, a seed or tuber from the transgenic plant, a paper sizing and/or
CC	coating composition comprising a glucan produced in a plant transformed
CC	with a gene encoding the mutant GTF, wild type or, starch, a latex,
CC	thermoplastic molecule or their combinations or glucan and starch where
CC	the glucan is produced in the amyloplast and/or vacuole or a maize line
CC	deficient in starch biosynthesis, transformed with a gene encoding a
CC	glucosyltransferase B or D enzyme, wild-type or mutant and a paper
CC	comprising the glucan (paper sizing/coating agent). The vector is useful
CC	for producing the glucan in a plant. The method comprises transforming a
CC	plant cell with the vector, growing the plant cell under plant growing
CC	conditions to produce a regenerated plant and inducing expression of the
CC	polynucleotide for a time sufficient to produce the glucan in the
CC	regenerated plant, where the vector contains a transit sequence from
CC	ribulose biphosphate carboxylase small subunit, waxy, brittle-1 and
CC	chlorophyll AB binding protein of potato or the vacuole of sugar beet.
CC	Glucans are useful as substitutes for and additions to modified starch
CC	and latexes in paper manufacture. Unlike prior art techniques, which
CC	require input materials that produce chemical effluents, paper
CC	manufacture utilizing the glucan produced by GTF, which utilizes
CC	biologically produced input materials, is more cost-effective and
CC	environmentally friendly. Moreover, glucans also exhibit thermoplastic
CC	properties and impart gloss to the paper during coating step.
CC	The present sequence represents a GFPD mutant of the invention.
CC	Note: The present sequence is not shown in the specification but
CC	was created by the indexer using the GFPD sequence appearing as AAU98029
CC	and the information in claim 36.
XX	

SQ Sequence 1430 AA;
 Query Match 66.4%; Score 75; DB 23; Length 1430;
 Best Local Similarity 75.0%; Pred. No. 0.007;
 Matches 15; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
 Oy 1 AIDHLSILEAMSGNDNDYK 20
 DB 495 AINHLSTLEAMSDNDPQYK 514
 RESULT 3
 AAU98042
 ID AAU98042 standard; Protein; 1430 AA.
 AC AAU98042;
 DT 27-AUG-2002 (first entry)
 DE S. mutans glucosyltransferase GTFD mutant T589E.
 XX Glucosyltransferase; GTFD; transgenic plant; paper sizing;
 KW coating composition; glucan; starch; latex; thermoplastic molecule;
 KW amyloplast; vacuole; paper manufacture; mutant; mutain.
 XX Streptococcus mutans.
 OS Synthetic.
 XX Key Location/Qualifiers
 FH Misc-difference 589
 FT /note= "Wild-type Thr substituted by Glu"
 FT US2002031826-A1.
 EN 14-MAR-2002.
 PD 19-DEC-2000; 2000US-0740274.
 PF 11-DEC-1998; 98US-0210361.
 PR 07-JUN-1995; 95US-0478704.
 PR 07-JUN-1995; 95US-0482711.
 PR 07-JUN-1995; 95US-0485243.
 PR 16-JAN-1998; 98US-0007999.
 PR 16-JAN-1998; 98US-0008172.
 PR 20-JAN-1998; 98US-0009620.
 PA (NICH/) NICHOLS S E.
 XX Nichols SE;
 PI WPI: 2002-414332/44.
 DR Glucosyltransferase B or D protein useful for producing a glucan useful
 PT as substitutes for and additions to modified starch and latexes in
 PT paper manufacture, comprises mutations in specific positions -
 XX Claim 36; Page -; 44pp; English.
 XX The invention an isolated protein comprising a glucosyltransferase
 CC (GTF) B polypeptide having changes at position from I448V, D457N,
 CC D567T, K1014T, D457N/D567T, D457N/D571K, D567T/D571K,
 CC D567T/D571K/K1014T, I448V/D457N/D567T/D571K/K779Q/K1014T,
 CC Y169A/Y110A/Y117A, and K779Q or a GTF D polypeptide having
 CC changes at positions from T589D, T589E, N471D, N471D/T589D, and
 CC N471D/T589E. Also included are a glucan produced by the GTF mutant,
 CC an isolated polynucleotide which encodes pi or p2, or its complementary
 CC polynucleotide, a ribonucleic acid sequence encoding the GTF mutant,
 CC an expression cassette comprising the polynucleotide operably linked to a
 CC promoter, a vector comprising the expression cassette, host cell
 CC introduced with the vector, a transgenic plant comprising the
 CC vector, a seed or tuber from the transgenic plant, a paper sizing and/or
 CC coating composition comprising a glucan produced in a plant transformed
 CC with a gene encoding the mutant GTF, wild type or, starch, a latex,

CC thermoplastic molecule or their combinations or glucan and starch where
 CC the glucan is produced in the amyloplast and/or vacuole or a maize line
 CC deficient in starch biosynthesis, transformed with a gene encoding a
 CC glucosyltransferase B or D enzyme, wild-type or mutant and a paper
 CC comprising the glucan (paper sizing/coating agent). The vector is useful
 CC for producing a glucan in a plant. The method comprises transforming a
 CC plant cell with the vector, growing the plant cell under plant growing
 CC conditions to produce a regenerated plant and inducing expression of the
 CC polynucleotide for a time sufficient to produce the glucan in the
 CC regenerated plant, where the vector contains a transit sequence from
 CC ribulose biphosphate carboxylase small subunit, waxy, brittle-1 and
 CC chlorophyll AB binding protein to produce a transgenic plant, and glucan
 CC is produced in the amyloplast of potato or the vacuole of sugar beet.
 CC Glucans are useful as substitutes for and additions to modified starch
 CC and latexes in paper manufacture. Unlike prior art techniques, which
 CC require input materials that produce chemical effluents, paper
 CC manufacture utilizing the glucan produced by GTF, which utilizes
 CC biologically produced input materials, is more cost-effective and
 CC environmentally friendly. Moreover, glucans also exhibit thermoplastic
 CC properties and impart gloss to the paper during coating step.
 CC The present sequence represents a GTFD mutant of the invention.
 CC Note: The present sequence is not shown in the specification but
 CC was created by the indexer using the GTFD sequence appearing as AAU98029
 CC and the information in claim 36.
 SQ Sequence 1430 AA;
 Query Match 66.4%; Score 75; DB 23; Length 1430;
 Best Local Similarity 75.0%; Pred. No. 0.007;
 Matches 15; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
 Oy 1 AIDHLSILEAMSGNDNDYK 20
 DB 495 AINHLSTLEAMSDNDPQYK 514
 RESULT 4
 AAU98043
 ID AAU98043 standard; Protein; 1430 AA.
 AC AAU98043;
 DT 27-AUG-2002 (first entry)
 DE S. mutans glucosyltransferase GTFD mutant N471D.
 XX Glucosyltransferase; GTFD; transgenic plant; paper sizing;
 KW coating composition; glucan; starch; latex; thermoplastic molecule;
 KW amyloplast; vacuole; paper manufacture; mutant; mutain.
 XX Streptococcus mutans.
 OS Synthetic.
 XX Key Location/Qualifiers
 FH Misc-difference 471
 FT /note= "Wild-type Asn substituted by Asp"
 FT US2002031826-A1.
 EN 14-MAR-2002.
 PD 19-DEC-2000; 2000US-0740274.
 PF 11-DEC-1998; 98US-0210361.
 PR 07-JUN-1995; 95US-0478704.
 PR 07-JUN-1995; 95US-0482711.
 PR 07-JUN-1995; 95US-0485243.
 PR 16-JAN-1998; 98US-0007999.
 PR 16-JAN-1998; 98US-0008172.
 PR 20-JAN-1998; 98US-0009620.
 PA (NICH/) NICHOLS S E.
 XX Nichols SE;
 PI WPI: 2002-414332/44.
 DR Glucosyltransferase B or D protein useful for producing a glucan useful
 PT as substitutes for and additions to modified starch and latexes in
 PT paper manufacture, comprises mutations in specific positions -
 XX Claim 36; Page -; 44pp; English.
 XX The invention an isolated protein comprising a glucosyltransferase
 CC (GTF) B polypeptide having changes at position from I448V, D457N,
 CC D567T, K1014T, D457N/D567T, D457N/D571K, D567T/D571K,
 CC D567T/D571K/K1014T, I448V/D457N/D567T/D571K/K779Q/K1014T,
 CC Y169A/Y110A/Y117A, and K779Q or a GTF D polypeptide having
 CC changes at positions from T589D, T589E, N471D, N471D/T589D, and
 CC N471D/T589E. Also included are a glucan produced by the GTF mutant,
 CC an isolated polynucleotide which encodes pi or p2, or its complementary
 CC polynucleotide, a ribonucleic acid sequence encoding the GTF mutant,
 CC an expression cassette comprising the polynucleotide operably linked to a
 CC promoter, a vector comprising the expression cassette, host cell
 CC introduced with the vector, a transgenic plant comprising the
 CC vector, a seed or tuber from the transgenic plant, a paper sizing and/or
 CC coating composition comprising a glucan produced in a plant transformed
 CC with a gene encoding the mutant GTF, wild type or, starch, a latex,

PI Nichols SE;
 XX WPI; 2002-414332/44.
 DR
 XX Glucosyltransferase B or D protein useful for producing a glucan useful
 PT as substrates for and additions to modified starch and latexes in
 XX paper manufacture, comprises mutations in specific positions -
 XX
 PS Claim 36; Page -; 44pp; English.
 XX
 XX The invention an isolated protein comprising a glucosyltransferase
 CC (GTF) B polypeptide having changes at position from 1448V, D457N,
 CC D567T, K1014T, D457N/D567T, D457N/D571K, D567T/D571K,
 CC D567T/D571K/K1014T, 1448V/D457N/D567T/D571K/K779Q/K1014T,
 CC Y169A/Y170A/Y171A, and K779Q or a GTF D polypeptide having
 CC changes at positions from 1589D, 1589E, N471D, N471D/1589D, and
 CC N471D/1589E. Also included are a glucan produced by the GTF mutant,
 CC an isolated polynucleotide which encodes P1 or P2, or its complementary
 CC polynucleotide, a ribonucleic acid sequence encoding the GTF mutant,
 CC an expression cassette comprising the polynucleotide operably linked to a
 CC promoter, a vector comprising the expression cassette, host cell
 CC introduced with the vector, a transgenic plant comprising the
 CC vector, a seed or tuber from the transgenic plant, a paper sizing and/or
 CC coating composition comprising a glucan produced in a plant transformed
 CC with a gene encoding the mutant GTF, wild type or, starch, a latex,
 CC thermoplastic molecule or their combinations or glucan and starch where
 CC the glucan is produced in the amyloplast and/or vacuole or a maize line
 CC deficient in starch biosynthesis, transformed with a gene encoding a
 CC glucosyltransferase B or D enzyme, wild-type or mutant and a paper
 CC comprising the glucan (paper sizing/coating agent). The vector is useful
 CC for producing a glucan in a plant. The method comprises transforming a
 CC plant cell with the vector, growing the plant cell under plant growing
 CC conditions to produce a regenerated plant and inducing expression of the
 CC polynucleotide for a time sufficient to produce the glucan in the
 CC regenerated plant, where the vector contains a transit sequence from
 CC ribulose biphosphate carboxylase small subunit, waxy, brittle-1 and
 CC chlorophyll AB binding protein to produce a transgenic plant, and glucan
 CC is produced in the amyloplast of potato or the vacuole of sugar beet.
 CC Glucans are useful as substitutes for and additions to modified starch
 CC and latexes in paper manufacture. Unlike prior art techniques, which
 CC require input materials that produce chemical effluents, paper
 CC manufacture utilizing the glucan produced by GTF, which utilizes
 CC biologically produced input materials, is more cost-effective and
 CC environmentally friendly. Moreover, glucans also exhibit thermoplastic
 CC properties and impart gloss to the paper during coating step.
 CC The present sequence represents a GTF mutant of the invention.
 CC Note: The present sequence is not shown in the specification but
 CC was created by the indexer using the GTF sequence appearing as AAU98029
 CC and the information in claim 36.
 XX
 XX Sequence 1430 AA;
 SQ
 Query Match 66.4%; Score 75; DB 23; Length 1430;
 Best Local Similarity 75.0%; Pred. No. 0.007;
 Matches 15; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
 QY 1 AINHLSTLEAMSGNDNDYVK 20
 Db 495 AINHLSTLEAMSGNDNDYVK 514
 RESULT 5
 ID AAU98044 standard; Protein; 1430 AA.
 AC
 XX AAU98044;
 XX
 DT 27-ATG-2002 (first entry)
 XX
 DE S. mutans [glucosyltransferase GTFD mutant N471D/1589D].
 XX
 KM Glucosyltransferase; GTFD; transgenic plant; paper sizing;
 KM coating composition; glucan; starch; latex; thermoplastic molecule;

KM amyloplast; vacuole; paper manufacture; mutant; mutein.
 XX
 XX Streptococcus mutans.
 OS
 OS Synthetic.
 XX
 XX Key Location/Qualifiers
 PH Misc-difference 471
 FT Misc-difference 589
 FT Misc-difference 589
 FT /note= "Wild-type Thr substituted by Asp"
 FT /note= "Wild-type Thr substituted by Asp"
 XX
 XX US2002031826-A1.
 XX
 XX 14-MAR-2002.
 XX
 XX 19-DEC-2000; 2000US-0740274.
 XX
 XX 11-DEC-1998; 98US-0210361.
 XX 07-JUN-1995; 95US-0478704.
 XX 07-JUN-1995; 95US-0482711.
 XX 07-JUN-1995; 95US-0485243.
 XX 16-JAN-1998; 98US-0007959.
 XX 16-JAN-1998; 98US-0008172.
 XX 20-JAN-1998; 98US-0009620.
 XX
 XX (NICH/) NICHOLS S E.
 XX
 XX Nichols SE;
 PI WPI; 2002-414332/44.
 DR
 XX Glucosyltransferase B or D protein useful for producing a glucan useful
 PT as substrates for and additions to modified starch and latexes in
 XX paper manufacture, comprises mutations in specific positions -
 XX
 XX Claim 36; Page -; 44pp; English.
 XX
 XX The invention an isolated protein comprising a glucosyltransferase
 CC (GTF) B polypeptide having changes at position from 1448V, D457N,
 CC D567T, K1014T, D457N/D567T, D457N/D571K, D567T/D571K,
 CC D567T/D571K/K1014T, 1448V/D457N/D567T/D571K/K779Q/K1014T,
 CC Y169A/Y170A/Y171A, and K779Q or a GTF D polypeptide having
 CC changes at positions from 1589D, 1589E, N471D, N471D/1589D, and
 CC N471D/1589E. Also included are a glucan produced by the GTF mutant,
 CC an isolated polynucleotide which encodes P1 or P2, or its complementary
 CC polynucleotide, a ribonucleic acid sequence encoding the GTF mutant,
 CC an expression cassette comprising the polynucleotide operably linked to a
 CC promoter, a vector comprising the expression cassette, host cell
 CC introduced with the vector, a transgenic plant comprising the
 CC vector, a seed or tuber from the transgenic plant, a paper sizing and/or
 CC coating composition comprising a glucan produced in a plant transformed
 CC with a gene encoding the mutant GTF, wild type or, starch, a latex,
 CC thermoplastic molecule or their combinations or glucan and starch where
 CC the glucan is produced in the amyloplast and/or vacuole or a maize line
 CC deficient in starch biosynthesis, transformed with a gene encoding a
 CC glucosyltransferase B or D enzyme, wild-type or mutant and a paper
 CC comprising the glucan (paper sizing/coating agent). The vector is useful
 CC for producing a glucan in a plant. The method comprises transforming a
 CC plant cell with the vector, growing the plant cell under plant growing
 CC conditions to produce a regenerated plant and inducing expression of the
 CC polynucleotide for a time sufficient to produce the glucan in the
 CC regenerated plant, where the vector contains a transit sequence from
 CC ribulose biphosphate carboxylase small subunit, waxy, brittle-1 and
 CC chlorophyll AB binding protein to produce a transgenic plant, and glucan
 CC is produced in the amyloplast of potato or the vacuole of sugar beet.
 CC Glucans are useful as substitutes for and additions to modified starch
 CC and latexes in paper manufacture. Unlike prior art techniques, which
 CC require input materials that produce chemical effluents, paper
 CC manufacture utilizing the glucan produced by GTF, which utilizes
 CC biologically produced input materials, is more cost-effective and
 CC environmentally friendly. Moreover, glucans also exhibit thermoplastic
 CC properties and impart gloss to the paper during coating step.
 CC The present sequence represents a GTF mutant of the invention.

CC Note: The present sequence is not shown in the specification but
 CC was created by the indexer using the GTFD sequence appearing as AAU98029
 CC and the information in claim 36.

XX Sequence 1430 AA;

Query Match 66.4%; Score 75; DB 23; Length 1430;
 Best Local Similarity 75.0%; Pred. No. 0.007;
 Matches 15; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 AIDHLSILEAMSGNDNDYK 20

DB 495 AINHLSTLEAMSDNDPQYK 514

RESULT 6
 AAU98045
 ID AAU98045 standard; Protein; 1430 AA.

XX AAU98045;

DT 27-AUG-2002 (first entry)

DE S. mutans glucosyltransferase GTFD mutant N471D/T589E.

XX Glucosyltransferase; GTFD; transgenic plant; paper sizing;
 KW coating composition; glucan; starch; latex; thermoplastic molecule;
 KM amyloplast; vacuole; paper manufacture; mutant; mutlein.

OS Streptococcus mutans.

XX Synthetic.

XX Key Location/Qualifiers

FT Misc-difference 471 /note= "Wild-type Asn substituted by Asp"

FT Misc-difference 589 /note= "Wild-type Thr substituted by Glu"

FT US2002031826-A1.

PD 14-MAR-2002.

PF 19-DEC-2000; 2000US-0740274.

XX 11-DEC-1998; 98US-0210361.

PR 07-JUN-1995; 95US-0478704.

PR 07-JUN-1995; 95US-0482711.

PR 07-JUN-1995; 95US-0485243.

PR 16-JAN-1998; 98US-0007999.

PR 16-JAN-1998; 98US-0008172.

PR 20-JAN-1998; 98US-0009620.

XX (NICH/) NICHOLS S E.

XX Nichols SE;

XX WPI; 2002-414332/44.

XX Glucosyltransferase B or D protein useful for producing a glucan useful

PT as substitutes for and additions to modified starch and latexes in

PT paper manufacture, comprises mutations in specific positions

XX Claim 36; Page -; 44pp; English.

XX The invention an isolated protein comprising a glucosyltransferase

CC an expression cassette comprising the polynucleotide operably linked to a
 CC promoter, a vector comprising the expression cassette, host cell
 CC introduced with the vector, a transgenic plant comprising the
 CC vector, a seed or tuber from the transgenic plant, a paper sizing and/or
 CC coating composition comprising a glucan produced in a plant transformed
 CC with a gene encoding the mutant GTF, wild type or, starch, a latex,
 CC thermoplastic molecule or their combinations or glucan and starch where
 CC the glucan is produced in the amyloplast and/or vacuole or a maize line
 CC deficient in starch biosynthesis, transformed with a gene encoding a
 CC glucosyltransferase B or D enzyme, wild-type or mutant and a paper
 CC comprising the glucan (paper sizing/coating agent). The vector is useful
 CC for producing a glucan in a plant. The method comprises transforming a
 CC plant cell with the vector, growing the plant cell under plant growing
 CC conditions to produce a regenerated plant and inducing expression of the
 CC polynucleotide for a time sufficient to produce the glucan in the
 CC regenerated plant, where the vector contains a transit sequence from
 CC ribulose biphosphate carboxylase small subunit, waxy, brittle-1 and
 CC chlorophyll AB binding protein to produce a transgenic plant, and glucan
 CC is produced in the amyloplast of potato or the vacuole of sugar beet.
 CC Glucans are useful as substitutes for and additions to modified starch
 CC and latexes in paper manufacture. Unlike prior art techniques, which
 CC require input materials that produce chemical effluents, paper
 CC manufacture utilizing the glucan produced by GTF, which utilizes
 CC biologically produced input materials, is more cost-effective and
 CC environmentally friendly. Moreover, glucans also exhibit thermoplastic
 CC properties and impart gloss to the paper during coating step.
 CC The present sequence represents a GTFD mutant of the invention.
 CC Note: The present sequence is not shown in the specification but
 CC was created by the indexer using the GTFD sequence appearing as AAU98029
 CC and the information in claim 36.

XX Sequence 1430 AA;

Query Match 66.4%; Score 75; DB 23; Length 1430;
 Best Local Similarity 75.0%; Pred. No. 0.007;
 Matches 15; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 AIDHLSILEAMSGNDNDYK 20

DB 495 AINHLSTLEAMSDNDPQYK 514

RESULT 7
 AAU79285
 ID AAU79285 standard; Protein; 1017 AA.

XX AAU79285;

DT 13-AUG-2002 (first entry)

DE Streptococcus mutans monoclonal antibody-related protein #2.

XX Antibody; dental caries; water insoluble glucan synthetase;

KM anti-carries; glucosyl transferase-B; immunotherapy.

OS Streptococcus mutans.

XX JP2002114709-A.

XX 16-APR-2002.

XX 04-OCT-2000; 2000JP-0304889.

XX 04-OCT-2000; 2000JP-0304889.

XX (UYNI-) UNIV NIPPON.

XX WPI; 2002-448101/48.

XX Anti-carries agent composed of a monoclonal antibody against an
 PT inhibitory enzyme against water insoluble glucan synthetase of glucosyl
 PT transferase-B (GTF-B) of Streptococcus mutans

PS Claim 4; Page 17-19; 28pp; Japanese.
 XX
 CC The invention relates to a monoclonal antibody against dental caries and
 CC an anti-caries agent composed of a monoclonal antibody produced by
 CC Streptococcus mutans, particularly mouse-hybridoma MHP126 (FERM P-17566)
 CC or mouse-hybridoma MHP136 (FERM P-17567), against an enzyme having
 CC inhibitive activity against water insoluble glucan synthetase of glucosyl
 CC transferase-B. The monoclonal antibody specifically inhibits water
 CC insoluble glucan synthetase of Streptococcus mutans produced glucosyl
 CC transferase-B and is used in the immunotherapy of dental caries. This
 CC sequence represents a Streptococcus mutans monoclonal antibody-related
 CC protein.
 XX
 SQ Sequence 1017 AA;
 Qy Query Match 64.6%; Score 73; DB 23; Length 1017;
 Db Best Local Similarity 73.7%; Pred. No. 0.0098;
 Matches 14; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
 1 AIDHLSILEAMSGNDNDYV 19
 447 ANDHLSILEAMSGNDNDYV 465
 RESULT 8
 AAU98027 ID AAU98027 standard; Protein; 1475 AA.
 AC AAU98027;
 XX
 DT 27-AUG-2002 (first entry)
 XX
 DE S. mutans glucosyltransferase GTFB.
 XX
 KM Glucosyltransferase; GTFB; transgenic plant; paper sizing;
 KM coating composition; glucan; starch; latex; thermoplastic molecule;
 KM amyloplast; vacuole; paper manufacture.
 XX
 OS Streptococcus mutans.
 OS
 PN US2002031826-A1.
 PN
 XX 14-MAR-2002.
 PD
 XX 19-DEC-2000; 2000US-0740274.
 PF
 XX 11-DEC-1998; 98US-0210361.
 PR 07-JUN-1995; 95US-0478704.
 PR 07-JUN-1995; 95US-0482711.
 PR 07-JUN-1995; 95US-0485243.
 PR 16-JAN-1998; 98US-0007999.
 PR 16-JAN-1998; 98US-0008172.
 PR 20-JAN-1998; 98US-0009620.
 XX
 XX (NICH/) NICHOLS S E.
 PA
 XX Nichols SE;
 PI
 XX WPI; 2002-414332/44.
 DR N-FSDB; ABR52938.
 DR
 XX Glucosyltransferase B or D protein useful for producing a glucan useful
 PT as substitutes for and additions to modified starch and latexes in
 PT paper manufacture, comprises mutations in specific positions -
 PT
 XX Disclosure; Page 21-25; 44pp; English.
 PS
 XX The invention an isolated protein comprising a glucosyltransferase
 CC (GTF) B polypeptide having changes at position from I448V, D457N,
 CC D567T, K1014T, D457N/D567T, D457N/D571K, D567T/D571K,
 CC D567T/D571K/K1014T, I448V/D457N/D567T/D571K/K779Q/K1014T,
 CC Y169A/Y170A/Y171A, and K779Q or a GTF D polypeptide having
 CC changes at positions from T589D, T589E, N471D, N471D/T589D, and

CC N471D/T589E. Also included are a glucan produced by the GTF mutant,
 CC an isolated polynucleotide which encodes P1 or P2, or its complementary
 CC polynucleotide, a ribonucleic acid sequence encoding the GTF mutant,
 CC an expression cassette comprising the polynucleotide operably linked to a
 CC promoter, a vector comprising the expression cassette, host cell
 CC introduced with the vector, a transgenic plant comprising the
 CC vector, a seed or tuber from the transgenic plant, a paper sizing and/or
 CC coating composition comprising a glucan produced in a plant transformed
 CC with a gene encoding the mutant GTF, wild type or, starch, a latex,
 CC thermoplastic molecule or their combinations or glucan and starch where
 CC the glucan is produced in the amyloplast and/or vacuole or a maize line
 CC deficient in starch biosynthesis, transformed with a gene encoding a
 CC glucosyltransferase B or D enzyme, wild-type or mutant and a paper
 CC comprising the glucan (paper sizing/coating agent). The vector is useful
 CC for producing a glucan in a plant. The method comprises transforming a
 CC plant cell with the vector, growing the plant cell under plant growing
 CC conditions to produce a regenerated plant and inducing expression of the
 CC polynucleotide for a time sufficient to produce the glucan in the
 CC regenerated plant, where the vector contains a transit sequence from
 CC ribulose biphosphate carboxylase small subunit, waxy, brittle-1 and
 CC chlorophyll AB binding protein to produce a transgenic plant, and glucan
 CC is produced in the amyloplast of potato or the vacuole of sugar beet.
 CC Glucans are useful as substitutes for and additions to modified starch
 CC and latexes in paper manufacture. Unlike prior art techniques, which
 CC require input materials that produce chemical effluents, paper
 CC manufacture utilizing the glucan produced by GTF, which utilizes
 CC biologically produced input materials, is more cost-effective and
 CC environmentally friendly. Moreover, glucans also exhibit thermoplastic
 CC properties and impart gloss to the paper during coating step.
 CC The present sequence represents GTFB.
 XX
 SQ Sequence 1475 AA;
 Qy Query Match 64.6%; Score 73; DB 23; Length 1475;
 Db Best Local Similarity 73.7%; Pred. No. 0.015;
 Matches 14; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
 1 AIDHLSILEAMSGNDNDYV 19
 481 ANDHLSILEAMSGNDNDYV 499
 RESULT 9
 AAU98030 ID AAU98030 standard; Protein; 1475 AA.
 XX
 AC AAU98030;
 XX
 DT 27-AUG-2002 (first entry)
 XX
 DE S. mutans glucosyltransferase GTFB mutant I448V.
 XX
 XX Glucosyltransferase; GTFB; transgenic plant; paper sizing;
 KM coating composition; glucan; starch; latex; thermoplastic molecule;
 KM amyloplast; vacuole; paper manufacture; mutant; muten.
 XX
 OS Streptococcus mutans.
 OS
 PN US2002031826-A1.
 PN
 XX 14-MAR-2002.
 PD
 XX 19-DEC-2000; 2000US-0740274.
 PF
 XX 11-DEC-1998; 98US-0210361.
 PR 07-JUN-1995; 95US-0478704.
 PR 07-JUN-1995; 95US-0482711.
 PR 07-JUN-1995; 95US-0485243.
 PR

PR 16-JAN-1998; 98US-0007999.
 PR 16-JAN-1998; 98US-0008172.
 PR 20-JAN-1998; 98US-0009620.
 XX
 PA (NICH/) NICHOLS S E.
 PI Nichols SE;
 XX
 DR WPI; 2002-414332/44.
 XX
 PT Glucosyltransferase B or D protein useful for producing a glucan useful
 PT as substitutes for and additions to modified starch and latexes in
 PT paper manufacture, comprises mutations in specific positions -
 XX
 PS Claim 36; Page -; 44pp; English.
 CC The invention an isolated protein comprising a glucosyltransferase
 CC (GTF) B polypeptide having changes at position from I448V, D457N,
 CC D567T, K1014T, D457N/D567T, D457N/D571K, D567T/D571K,
 CC D567T/D571K/K1014T, I448V/D457N/D567T/K779Q/K1014T,
 CC Y169A/Y170A/Y171A, and K779Q or a GTF D polypeptide having
 CC changes at positions from T589D, T589E, N471D, N471D/T589D, and
 CC N471D/T589E. Also included are a glucan produced by the GTF mutant,
 CC an isolated polynucleotide which encodes P1 or P2, or its complementary
 CC polynucleotide, a ribonucleic acid sequence encoding the GTF mutant,
 CC an expression cassette comprising the polynucleotide operably linked to a
 CC promoter, a vector comprising the expression cassette, host cell
 CC introduced with the vector, a transgenic plant comprising the
 CC vector, a seed or tuber from the transgenic plant, a paper sizing and/or
 CC coating composition comprising a glucan produced in a plant transformed
 CC with a gene encoding the mutant GTF, wild type or, starch, a latex,
 CC thermoplastic molecule or their combinations or glucan and starch where
 CC the glucan is produced in the amyloplast and/or vacuole or a maize line
 CC deficient in starch biosynthesis, transformed with a gene encoding a
 CC glucosyltransferase B or D enzyme, wild-type or mutant and a paper
 CC comprising the glucan (paper sizing/coating agent). The vector is useful
 CC for producing a glucan in a plant. The method comprises transforming a
 CC plant cell with the vector, growing the plant cell under plant growing
 CC conditions to produce a regenerated plant and inducing expression of the
 CC polynucleotide for a time sufficient to produce the glucan in the
 CC regenerated plant, where the vector contains a transit sequence from
 CC ribulose biphosphate carboxylase small subunit, waxy, brittle-1 and
 CC chlorophyll AB binding protein to produce a transgenic plant, and glucan
 CC is produced in the amyloplast of potato or the vacuole of sugar beet.
 CC Glucans are useful as substitutes for and additions to modified starch
 CC and latexes in paper manufacture. Unlike prior art techniques, which
 CC require input materials that produce chemical effluents, paper
 CC manufacture utilizing the glucan produced by GTF, which utilizes
 CC biologically produced input materials, is more cost-effective and
 CC environmentally friendly. Moreover, glucans also exhibit thermoplastic
 CC properties and impart gloss to the paper during coating step.
 CC The present sequence represents a GTFB mutant of the invention.
 CC Note: The present sequence is not shown in the specification but
 CC was created by the indexer using the GTFB sequence appearing as AAU98027
 CC and the information in claim 36.
 XX
 SQ Sequence 1475 AA;
 Query Match 64.6%; Score 73; DB 23; Length 1475;
 Best Local Similarity 73.7%; Pred. NO. 0.015; Indels 0; Gaps 0;
 Matches 14; Conservative 1; Mismatches 4;
 QY 1 AIDHSTLEAMSGNDNDYV 19
 Db 481 ANDHSTLEAMSGNDNDYVYL 499
 RESULT 10
 AAU98031
 ID AAU98031 standard; Protein; 1475 AA.
 XX
 AC AAU98031;
 XX

DT 27-AUG-2002 (first entry)
 XX
 DE S. mutans glucosyltransferase GTFB mutant D457N.
 XX
 KW Glucosyltransferase; GTFB; transgenic plant; paper sizing;
 KW coating composition; glucan; starch; latex; thermoplastic molecule;
 KW amyloplast; vacuole; paper manufacture; mutant; muten.
 XX
 OS Streptococcus mutans.
 OS Synthetic.
 XX
 FH Key Location/Qualifiers
 FT Misc-difference 457
 PT /note= "Wild-type Asp substituted by Asn"
 XX
 PN US2002031826-A1.
 XX
 PD 14-MAR-2002.
 XX
 PE 19-DEC-2000; 2000US-0740274.
 XX
 PR 11-DEC-1998; 98US-0210361.
 PR 07-JUN-1995; 95US-0478704.
 PR 07-JUN-1995; 95US-0482711.
 PR 07-JUN-1995; 95US-0485243.
 PR 16-JAN-1998; 98US-0007999.
 PR 16-JAN-1998; 98US-0008172.
 PR 20-JAN-1998; 98US-0009620.
 XX
 PA (NICH/) NICHOLS S E.
 PI Nichols SE;
 XX
 DR WPI; 2002-414332/44.
 XX
 PT Glucosyltransferase B or D protein useful for producing a glucan useful
 PT as substitutes for and additions to modified starch and latexes in
 PT paper manufacture, comprises mutations in specific positions -
 XX
 PS Claim 36; Page -; 44pp; English.
 CC The invention an isolated protein comprising a glucosyltransferase
 CC (GTF) B polypeptide having changes at position from I448V, D457N,
 CC D567T, K1014T, D457N/D567T, D457N/D571K, D567T/D571K,
 CC D567T/D571K/K1014T, I448V/D457N/D567T/D571K/K779Q/K1014T,
 CC Y169A/Y170A/Y171A, and K779Q or a GTF D polypeptide having
 CC changes at positions from T589D, T589E, N471D, N471D/T589D, and
 CC N471D/T589E. Also included are a glucan produced by the GTF mutant,
 CC an isolated polynucleotide which encodes P1 or P2, or its complementary
 CC polynucleotide, a ribonucleic acid sequence encoding the GTF mutant,
 CC an expression cassette comprising the polynucleotide operably linked to a
 CC promoter, a vector comprising the expression cassette, host cell
 CC introduced with the vector, a transgenic plant comprising the
 CC vector, a seed or tuber from the transgenic plant, a paper sizing and/or
 CC coating composition comprising a glucan produced in a plant transformed
 CC with a gene encoding the mutant GTF, wild type or, starch, a latex,
 CC thermoplastic molecule or their combinations or glucan and starch where
 CC the glucan is produced in the amyloplast and/or vacuole or a maize line
 CC deficient in starch biosynthesis, transformed with a gene encoding a
 CC glucosyltransferase B or D enzyme, wild-type or mutant and a paper
 CC comprising the glucan (paper sizing/coating agent). The vector is useful
 CC for producing a glucan in a plant. The method comprises transforming a
 CC plant cell with the vector, growing the plant cell under plant growing
 CC conditions to produce a regenerated plant and inducing expression of the
 CC polynucleotide for a time sufficient to produce the glucan in the
 CC regenerated plant, where the vector contains a transit sequence from
 CC ribulose biphosphate carboxylase small subunit, waxy, brittle-1 and
 CC chlorophyll AB binding protein to produce a transgenic plant, and glucan
 CC is produced in the amyloplast of potato or the vacuole of sugar beet.
 CC Glucans are useful as substitutes for and additions to modified starch
 CC and latexes in paper manufacture. Unlike prior art techniques, which
 CC require input materials that produce chemical effluents, paper
 CC manufacture utilizing the glucan produced by GTF, which utilizes

CC biologically produced input materials, is more cost-effective and
CC environmentally friendly. Moreover, glucans also exhibit thermoplastic
CC properties and impart gloss to the paper during coating step.
CC The present sequence represents a GFPB mutant of the invention.
CC Note: The present sequence is not shown in the specification but
CC was created by the indexer using the GFPB sequence appearing as AAU98027
CC and the information in claim 36.
XX
SQ Sequence 1475 AA;
Query Match 64.6%; Score 73; DB 23; Length 1475;
Best Local Similarity 73.7%; Pred. No. 0.015;
Matches 14; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
QY 1 AIDHSLTEAWSGNDNDYV 19
DB 481 ANDHSLTEAWSGNDNDYV 499
RESULT 11
AAU98032 standard; Protein; 1475 AA.
XX
AC AAU98032;
XX
DT 27-AUG-2002 (first entry)
XX
DE S. mutans glucosyltransferase GFPB mutant D567T.
XX
KW Glucosyltransferase; GFPB; transgenic plant; paper sizing;
KW coating composition; glucan; starch; latex; thermoplastic molecule;
KW amyloplast; vacuole; paper manufacture; mutant; mutein.
XX
OS Streptococcus mutans.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Misc-difference 567 /note= "Wild-type Asp substituted by Thr"
XX
PN US2002031826-A1.
XX
PD 14-MAR-2002.
XX
PF 19-DEC-2000; 2000US-0740274.
XX
PR 11-DEC-1998; 98US-0210361.
PR 07-JUN-1995; 95US-0478704.
PR 07-JUN-1995; 95US-0482711.
PR 07-JUN-1995; 95US-0485243.
PR 16-JAN-1998; 98US-0007999.
PR 16-JAN-1998; 98US-0008172.
PR 20-JAN-1998; 98US-0009620.
XX
XX (NICH/) NICHOLS S E.
XX PI Nichols SE;
XX
DR WPI; 2002-414332/44.
XX
PT Glucosyltransferase B or D protein useful for producing a glucan useful
PT as substitutes for and additions to modified starch and latexes in
PT paper manufacture, comprises mutations in specific positions -
XX
PS Claim 36; Page -; 44p; English.
XX
CC The invention an isolated protein comprising a glucosyltransferase
CC (GTF) B polypeptide having changes at position from 1448V, D457N,
CC D567T, K1014T, D457N/D567T, D457N/D571K, D567T/D571K,
CC D567T/D571K/K1014T, 1448V/D457N/D567T/D571K/K1014T,
CC Y169A/Y170A/Y171A, and K779Q or a GFP D polypeptide having
CC changes at positions from T589D, T589E, N471D, N471D/T589D, and
CC N471D/T589E. Also included are a glucan produced by the GFP mutant,

CC an isolated polynucleotide which encodes P1 or P2, or its complementary
CC polynucleotide, a ribonucleic acid sequence encoding the GFP mutant,
CC an expression cassette comprising the polynucleotide operably linked to a
CC promoter, a vector comprising the expression cassette, host cell
CC introduced with the vector, a transgenic plant comprising the
CC vector, a seed or tuber from the transgenic plant, a paper sizing and/or
CC coating composition comprising a glucan produced in a plant transformed
CC with a gene encoding the mutant GFP, wild type or, starch, a latex,
CC thermoplastic molecule or their combinations or glucan and starch where
CC the glucan is produced in the amyloplast and/or vacuole or a maize line
CC deficient in starch biosynthesis, transformed with a gene encoding a
CC glucosyltransferase B or D enzyme, wild-type or mutant and a paper
CC comprising the glucan (paper sizing/coating agent). The vector is useful
CC for producing a glucan in a plant. The method comprises transforming a
CC plant cell with the vector, growing the plant cell under plant growing
CC conditions to produce a regenerated plant and inducing expression of the
CC polynucleotide for a time sufficient to produce the glucan in the
CC regenerated plant, where the vector contains a transit sequence from
CC ribulose biphosphate carboxylase small subunit, waxy, brittle-1 and
CC chlorophyll AB binding protein to produce a transgenic plant, and glucan
CC is produced in the amyloplast of potato or the vacuole of sugar beet.
CC Glucans are useful as substitutes for and additions to modified starch
CC and latexes in paper manufacture. Unlike prior art techniques, which
CC require input materials that produce chemical effluents, paper
CC manufacture utilizing the glucan produced by GFP, which utilizes
CC biologically produced input materials, is more cost-effective and
CC environmentally friendly. Moreover, glucans also exhibit thermoplastic
CC properties and impart gloss to the paper during coating step.
CC The present sequence represents a GFPB mutant of the invention.
CC Note: The present sequence is not shown in the specification but
CC was created by the indexer using the GFPB sequence appearing as AAU98027
CC and the information in claim 36.
XX
SQ Sequence 1475 AA;
Query Match 64.6%; Score 73; DB 23; Length 1475;
Best Local Similarity 73.7%; Pred. No. 0.015;
Matches 14; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
QY 1 AIDHSLTEAWSGNDNDYV 19
DB 481 ANDHSLTEAWSGNDNDYV 499

RESULT 12
AAU98033 standard; Protein; 1475 AA.
XX
AC AAU98033;
XX
DT 27-AUG-2002 (first entry)
XX
DE S. mutans glucosyltransferase GFPB mutant K1014T.
XX
KW Glucosyltransferase; GFPB; transgenic plant; paper sizing;
KW coating composition; glucan; starch; latex; thermoplastic molecule;
KW amyloplast; vacuole; paper manufacture; mutant; mutein.
XX
OS Streptococcus mutans.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Misc-difference 1014 /note= "Wild-type Iys substituted by Thr"
XX
PN US2002031826-A1.
XX
PD 14-MAR-2002.
XX
PF 19-DEC-2000; 2000US-0740274.
XX
PR 11-DEC-1998; 98US-0210361.
PR 07-JUN-1995; 95US-0478704.

CC Glucans are useful as substitutes for and additions to modified starch
CC and latexes in paper manufacture. Unlike prior art techniques, which
CC require input materials that produce chemical effluents, paper
CC manufacture utilizing the glucan produced by GRF, which utilizes
CC biologically produced input materials, is more cost-effective and
CC environmentally friendly. Moreover, glucans also exhibit thermoplastic
CC properties and impart gloss to the paper during coating step.
CC The present sequence represents a GRF mutant of the invention.
CC Note: The present sequence is not shown in the specification but
CC was created by the indexer using the GRF sequence appearing as AM98027
CC and the information in Claim 36.

SQ Sequence 1475 AA;

Query Match	64.6%	Score 73	DB 23	Length 1475
Best Local Similarity	73.7%	Pred. No. 0.015		
Matches 14	Conservative 1	Mismatches 4	Indels 0	Gaps 0

```
QY      1 AIDHSLLEAWSGNDNDYV 19  
         |||||  
Db     481 ANDHSLLEAWSNDNDTPYL 499
```

RESULT 14

ID AAU98035 standard; Protein; 1475 AA.

AC AAU98035;

DT 27-AUG-2002 (first entry)

DE S. mutans glucosyltransferase GTFB mutant D457N/D571K.

KM Glucosyltransferase; GTFB; transgenic plant; paper sizing;
KM coating composition: glucan, starch; latex; thermoplastic molecule
KM amyloplast; vacuole; paper manufacture; mutant; mutain.

OS Streptococcus mutans,
OS synthetic.

Key	Location/Qualifiers
FH	

FT	/note= "Wild-type Asp substituted by Asn"
----	---

FT	/note= "Wild-type Asp substituted by Lys"
----	---

PN US2002031826-A1

PD 14-MAR-2002

19-DEC-2000; 2000US-0740274.

PR 11-DEC-1998; 98US-0210361.

PR 07-JUN-1995; 95US-0482711.

PR 16-JAN-1998; 9BUS-0007999

PR 20-JAN-1998; 98US-0009620

PA (NICH/) NICHOLS S E.

PI Nichols SE;

DR WPI; 2002-414332/44.

PT Glucosyltra

PT paper manufacture, c

PS Claim 36; Page -; 44pp; English.

CC The invention an isolated protein comprising a glucosyltransferase

CC (GTF) B polypeptide having changes at position from I448V, D457N
CC K1014T, D457N/D567T, D457N/D571K, D567E/D571K,
D567T, K1014T, D457N/D567T, D457N/D571K,

CC Glucans are useful as substitutes for and additions to modified starch
CC and latexes in paper manufacture. Unlike prior art techniques, which
CC require input materials that produce chemical effluents, paper
CC manufacture utilizing the glucan produced by GRF, which utilizes
CC biologically produced input materials, is more cost-effective and
CC environmentally friendly. Moreover, glucans also exhibit thermoplastic
CC properties and impart gloss to the paper during coating step.
CC The present sequence represents a GRF mutant of the invention.
CC Note: The present sequence is not shown in the specification but
CC was created by the indexer using the GRF sequence appearing as AA09802
CC and the information in claim 36.

sq Sequence 1475 AA;

Query Match	64.6%;	Score 73;	DB 23;	Length 1475;
Best Local Similarity	73.7%;	Pred. No. 0.015;		
Matches 14;	Conservative 1;	Mismatches 4;	Indels 0;	Gaps 0

```

QY      1 AIDHLSILEAWSGNDNDYV 19
         | | | | | | | | | | :
Db      481 ANDHLSILEAWSDNDTPYL 499

```

RESULT 15

ID AAU98036 standard; Protein; 1475 AA

AC AAU98036;

DT 27-AUG-2002 (first entry)

DE S. mutans glucosyltransferase GTFB mutant D567T/D571K.

KM Glucosyltransferase; GTFB; transgenic plant, paper sizing;
KM coating composition; glucan; starch; latex; thermoplastic molecule;
KM amyloplast; vacuole; paper manufacture; mutant; mutein.

OS Streptococcus mutans.

XX

Misc-difference	567
FT	

FT Misc-difference 571

XX

PN US2002031826-A1.
XX
PD 14-MAR-2002.
XX
PF 19-DEC-2000; 2000US-0740274.
XX
PR 11-DEC-1998; 98US-0210361.
PR 07-JUN-1995; 95US-0478704.
PR 07-JUN-1995; 95US-0482711.
PR 16-JAN-1998; 98US-0007999.
PR 16-JAN-1998; 98US-0008172.
PR 20-JAN-1998; 98US-0009620.
XX
PA (NICH/) NICHOLS S E.
XX
PI Nichols SE;
XX
DR WPI; 2002-414332/44.
XX
XX

Glucosyltransferase B or D protein useful for producing a glucan useful as substitutes for and additions to modified starch and latexes in paper manufacture, comprises mutations in specific positions -
Claim 36; Page -: 44pp; English.

The invention an isolated protein comprising a glucosyltransferase (GTF) B polypeptide having changes at position from 1448V, D457N, D567T, K1014T, D457N/D567T, D457N/D571K, D567T/D571K, D567T/D571K/K1014T, I448V/D457N/D567T/D571K/K779Q/K1014T, Y169A/Y170A/Y171A, and K779Q or a GTF D polypeptide having changes at positions from T589D, T589E, N471D, N471D/T589D, and N471D/T589E. Also included are a glucan produced by the GTF mutant, an isolated polynucleotide which encodes P1 or P2, or its complementary polynucleotide, a ribonucleic acid sequence encoding the GTF mutant, an expression cassette comprising the polynucleotide operably linked to a promoter, a vector comprising the expression cassette, host cell introduced with the vector, a transgenic plant comprising the vector, a seed or tuber from the transgenic plant, a paper sizing and/or coating composition comprising a glucan produced in a plant transformed with a gene encoding the mutant GTF, wild type or, starch, a latex, thermoplastic molecule or their combinations or glucan and starch where the glucan is produced in the amyloplast and/or vacuole or a maize line deficient in starch biosynthesis, transformed with a gene encoding a glucosyltransferase B or D enzyme, wild-type or mutant and a paper comprising the glucan (paper sizing/coating agent). The vector is useful for producing a glucan in a plant. The method comprises transforming a plant cell with the vector, growing the plant cell under plant growing conditions to produce a regenerated plant and inducing expression of the polynucleotide for a time sufficient to produce the glucan in the regenerated plant, where the vector contains a transit sequence from ribulose biphosphate carboxylase small subunit, waxy, brittle-1 and chlorophyll AB binding protein to produce a transgenic plant, and glucan is produced in the amyloplast of potato or the vacuole of sugar beet. Glucans are useful as substitutes for and additions to modified starch and latexes in paper manufacture. Unlike prior art techniques, which require input materials that produce chemical effluents, paper manufacture utilizing the glucan produced by GTF, which utilizes biologically produced input materials, is more cost-effective and environmentally friendly. Moreover, glucans also exhibit thermoplastic properties and impart gloss to the paper during coating step. CC The present sequence represents a GTFB mutant of the invention. CC Note: The present sequence is not shown in the specification but CC was created by the indexer using the GTFB sequence appearing as AA098027 CC and the information in claim 36.
XX
SQ Sequence 1475 AA;

Query Match 64.6%; Score 73; DB 23; Length 1475;
Best Local Similarity 73.7%; Pred. No. 0.015;
Matches 14; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 AIDHLSILEAWSGNDNDYV 19

Db 481 AIDHLSILEAWSGNDNDYV 499
Search completed: November 13, 2003, 09:38:27
Job time : 38.5166 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 13, 2003, 09:31:40 ; Search time 13.0379 Seconds
(without alignments)
154.898 Million cell updates/sec

Title: US-09-290-049A-13
Perfect score: 113
Sequence: 1 AIDHSLTEAMSGNDNDYKQ 21

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	108	95.6	1365	2	A41483	glucosyltransferase
2	75	66.4	1431	2	A45866	dextranucrase (EC
3	73	64.6	1475	2	B31335	glcF protein precu
4	73	64.6	1577	2	T30858	glucosyltransferase
5	71	62.8	1375	2	T30345	dextranucrase (EC
6	69	61.1	1508	2	T31098	probable dextran
7	69	61.1	1518	2	A44811	glucosyltransferase
8	67	59.3	1449	2	T30857	glucosyltransferase
9	67	59.3	1449	2	T30552	glucosyltransferase
10	62	54.9	1592	2	A38175	glucosyltransferase
11	61	54.0	1599	2	S22737	glucosyltransferase
12	49	43.4	232	2	B64176	hypothetical prote
13	49	43.4	1428	1	ISB172	DNA topoisomerase
14	48	42.5	366	2	AB3211	hypothetical prote
15	48	42.5	378	2	T04254	hypothetical prote
16	48	42.5	695	2	B72129	probable outer mem
17	48	42.5	695	2	F66493	probable outer mem
18	48	42.5	704	2	C81542	conserved hypothet
19	48	42.5	3206	1	GMS3PV	genome polyprotein
20	47	41.6	298	2	T15906	hypothetical prote
21	47	41.6	680	2	T39858	hypothetical prote
22	47	41.6	769	2	D86335	T20H2.6 protein -
23	46	40.7	287	2	T36413	probable ABC-type
24	46	40.7	458	2	H86314	hypothetical prote
25	46	40.7	514	2	F87592	hypothetical prote
26	46	40.7	536	2	T13261	hypothetical prote
27	46	40.7	1252	2	H97178	ATP-dependent exon
28	46	40.7	1802	2	T00020	bacterial blight-r
29	45	39.8	2366	2	S10317	toxin B - Clostrid

30	45	39.8	2367	2 S70172	toxin B - Clostrid
31	44	38.9	342	2 E69581	acetoin dehydrogen
32	44	38.9	389	2 T17601	hypothetical prote
33	44	38.9	468	2 S61964	probable membrane
34	44	38.9	588	2 C83836	subtilisin-type pr
35	43.5	38.5	356	2 T48354	hypothetical prote
36	43	38.1	76	2 T42309	hypothetical prote
37	43	38.1	162	2 C97754	hypothetical prote
38	43	38.1	177	2 T40825	hypothetical prote
39	43	38.1	221	2 A67411	DNA-binding respon
40	43	38.1	270	2 AC2987	conserved hypothet
41	43	38.1	270	2 E98296	hypothetical prote
42	43	38.1	272	2 H72596	hypothetical prote
43	43	38.1	424	2 AD2238	hypothetical prote
44	43	38.1	429	2 T45040	hypothetical prote
45	43	38.1	697	2 H71525	probable outer mem

ALIGNMENTS

RESULT 1

A41483
Glucosyltransferase (EC 2.4.1.1) glfs precursor - Streptococcus sobrinus
C:Species: Streptococcus sobrinus
C:Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 15-Oct-1999
C:Accession: A41483
R:Gilmore, K.S.; Russell, R.R.B.; Ferretti, J.J.
I:Infect. Immun. 58, 2452-2458, 1990
A>Title: Analysis of the Streptococcus downei glfs gene, which specifies a glucosyltransferase
A:Reference number: A41483; MUID:90316665; PMID:2142479
A:Accession: A41483
A:Molecule type: DNA
A:Residues: 1-1365 <GIL>
A:Cross-references: GB:M30943; NID:G153652; PIDN:AAA26898.1; PID:G153653
C:Genetics:
A:Gene: glfs
C:Superfamily: cpl repeat homology
C:Keywords: glucosyltransferase; hexosyltransferase

Query Match 95.6%; Score 108; DB 2; Length 1365;
Best local similarity 100.0%; Pred. No. 1.3e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AIDHSLTEAMSGNDNDYK 20
DB 467 AIDHSLTEAMSGNDNDYK 486

RESULT 2

A45866
dextranucrase (EC 2.4.1.5) precursor - Streptococcus mutans
C:Species: Streptococcus mutans
C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 07-May-1999
C:Accession: A45866
R:Honda, O.; Kato, C.; Kuramitsu, H.K.
J:Gen. Microbiol. 136, 2099-2105, 1990
A>Title: Nucleotide sequence of the Streptococcus mutans glfD gene encoding the glucosyltransferase
A:Reference number: A45866; MUID:9110958; PMID:2146600
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1431 <HON>
A:Cross-references: GB:M29296
C:Superfamily: cpl repeat homology
C:Keywords: glucosyltransferase; hexosyltransferase
F:181-201/Domain: cpl repeat homology <CP1>
F:1127-1146/Domain: cpl repeat homology <CP2>
F:1192-1211/Domain: cpl repeat homology <CP3>
F:1257-1276/Domain: cpl repeat homology <CP4>
F:1277-1297/Domain: cpl repeat homology <CP5>
F:1321-1340/Domain: cpl repeat homology <CP6>
F:1341-1361/Domain: cpl repeat homology <CP6>

F:1385-1404/Domain: cpl repeat homology <CP7>

Query Match 66.4%; Score 75; DB 2; Length 1431;
 Best Local Similarity 75.0%; Pred. No. 0.0023;
 Matches 15; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 AIDHLSILEAMSGNDNDYK 20
 ||:|||||:|||||:|
 Db 495 AINHLSILEAMSGNDNDYK 514

RESULT 3

gtrB protein precursor - Streptococcus mutans
 C:Species: Streptococcus mutans
 C:Date: 23-Oct-1990 #sequence_revision 23-Oct-1990 #text_change 15-Oct-1999

C:Accession: B33135; A33128
 R:Shiroza, T.; Ueda, S.; Kuramitsu, H.K.
 J. Bacteriol. 169, 4263-4270, 1987

A:Title: Sequence analysis of the gtrB gene from Streptococcus mutans.
 A:Reference number: A33135; MUID:87308013; PMID:3040685

A:Accession: B33135
 A:Status: preliminary
 A:Molecule type: DNA

A:Residues: 1-1475 <SH1>
 A:Cross-references: GB:M17361; NID:G153639; PIDN:AAA88588.1; PID:G153640
 R:Shiroza, T.; Ueda, S.; Kuramitsu, H.K.

submitted to the Protein Sequence Database, September 1990
 A:Reference number: A33128
 A:Accession: A33128
 A:Status: preliminary; not compared with conceptual translation
 A:Molecule type: DNA

A:Residues: 1-171,173-641,'N',643-1475 <SH2>
 A:Experimental source: strain GS-5

C:Superfamily: cpl repeat homology
 F:1096-1115/Domain: cpl repeat homology <CP1>
 F:1224-1243/Domain: cpl repeat homology <CP2>
 F:1289-1308/Domain: cpl repeat homology <CP3>
 F:1354-1373/Domain: cpl repeat homology <CP4>
 F:1419-1438/Domain: cpl repeat homology <CP5>

Query Match 64.6%; Score 73; DB 2; Length 1475;
 Best Local Similarity 73.7%; Pred. No. 0.0049;
 Matches 14; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 AIDHLSILEAMSGNDNDYK 19
 ||:|||||:|||||:|
 Db 481 AIDHLSILEAMSGNDNDYK 499

RESULT 4
 T30858
 glucosyltransferase - Streptococcus salivarius
 C:Species: Streptococcus salivarius
 C:Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 22-Oct-1999

C:Accession: T30858
 R:Simpson, C.L.; Giffard, P.M.; Jacques, N.A.
 Infect. Immun. 63, 609-621, 1995

A:Title: Streptococcus salivarius ATCC 25975 possesses at least two genes coding for pti
 A:Reference number: Z20909; MUID:95122197; PMID:7822030

A:Accession: T30858
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA

A:Residues: 1-1577 <SIM>
 A:Cross-references: EMBL:L55928; NID:G662380; PID:G662381; PIDN:AA041413.1
 C:Genetics: gtfC

Query Match 64.6%; Score 73; DB 2; Length 1577;
 Best Local Similarity 75.0%; Pred. No. 0.0053;
 Matches 15; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 AIDHLSILEAMSGNDNDYK 20

Db 588 AIAHLSILEAMSYNDHQYNK 607
 ||:|||||:|||||:|

RESULT 5

UT0345
 dextranucrase (EC 2.4.1.5) precursor - Streptococcus mutans (strain GS-5)
 N:Alternate names: sucrose 6-glucosyltransferase
 C:Species: Streptococcus mutans
 C:Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 24-Sep-1999

C:Accession: UT0345; C33135
 R:Ueda, S.; Shiroza, T.; Kuramitsu, H.K.
 Gene 69, 101-109, 1988

A:Title: Sequence analysis of the gtfC gene from Streptococcus mutans GS-5.
 A:Reference number: UT0345; MUID:89137980; PMID:2976010

A:Accession: UT0345
 A:Molecule type: DNA
 A:Residues: 1-1375 <UED>

A:Experimental source: GS-5
 R:Shiroza, T.; Ueda, S.; Kuramitsu, H.K.
 J. Bacteriol. 169, 4263-4270, 1987

A:Title: Sequence analysis of the gtfB gene from Streptococcus mutans.
 A:Reference number: A33135; MUID:87308013; PMID:3040685

A:Accession: C33135
 A:Status: preliminary
 A:Molecule type: DNA

A:Residues: 1-349 <SH1>
 A:Cross-references: GB:M17361
 C:Genetics: gtfC

C:Function: catalyzes the synthesis of both water-soluble and water-insoluble glucans
 A:Description: cpl repeat homology

C:Superfamily: cpl repeat homology
 C:Keywords: duplication; signal sequence #status predicted <SIG>
 F:1-34/Domain: signal sequence #status predicted <MAT>
 F:35-137/Product: glucosyltransferase #status predicted <MAT>

F:1126-1145/Domain: cpl repeat homology <CP1>
 F:1125-1272/Domain: cpl repeat homology <CP2>
 F:1318-1337/Domain: cpl repeat homology <CP3>

Query Match 62.8%; Score 71; DB 2; Length 1375;
 Best Local Similarity 73.7%; Pred. No. 0.0094;
 Matches 14; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 AIDHLSILEAMSGNDNDYK 19
 ||:|||||:|||||:|
 Db 507 AIDHLSILEAMSGNDNDYK 525

RESULT 6

T31098
 probable dextranucrase (EC 2.4.1.5), extracellular - Leuconostoc mesenteroides
 C:Species: Leuconostoc mesenteroides
 C:Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 11-May-2000

C:Accession: T31098
 R:Roncholski, V.; Renaud-Simeon, M.; Monahan, P.; Willemot, R.M.
 FEBS Microbiol. Lett. 159, 307-315, 1998

A:Title: Cloning and sequencing of a gene coding for an extracellular dextranucrase (DS
 A:Reference number: Z20981; MUID:98164374; PMID:9503626

A:Accession: T31098
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA

A:Residues: 1-1508 <MON>
 A:Cross-references: EMBL:AF030129; NID:G2766611; PID:G2766612; PIDN:AA095453.1
 A:Experimental source: strain NRRL B-1299

C:Genetics: dextr
 A:Function:

A:Description: produces dextran composed only of alpha(1-6) glucosidic bonds
 C:Keywords: glucosyltransferase, hexosyltransferase

Query Match 61.1%; Score 69; DB 2; Length 1508;
 Best Local Similarity 70.0%; Pred. No. 0.022;

A;Molecule type: DNA
A;Residues: 1-51 <GIF>

A:Cross-references: EMBL:Z11873
 C:Genetics:
 A:Gene: glfX
 C:Superfamily: cpl repeat homology
 C:Keywords: glycosyltransferase; hexosyltransferase
 P:1456-1475/domain: cpl repeat homology <CPR>

Query Match 54.0%; Score 61; DB 2; Length 1599;
 Best Local Similarity 57.1%; Pred. No. 0.43;
 Matches 12; Conservative 4; Mismatches 5; Indels 0; Gaps 0;
 QY 1 AIDHLSLEAMSGNDNDYVQ 21
 DB 491 ALANISLEAMSHNDPYVNE 511

RESULT 12
 Hypothetical protein Vng0156c [imported] - Halobacterium sp. NRC-1
 C:Species: Halobacterium sp. NRC-1
 C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001
 C:Accession: B84176
 R:Ng, W.V.; Kennedy, S.P.; Mahlars, G.G.; Bergquist, B.; Pan, M.; Shukla, H.D.; Laeky, S.; Jaitnauer, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jaidic Jung, K.H.; Alam, M.; Freitas, T.; 12176-12181, 2000
 Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000
 A:Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ehardt, H.; Lowe, T.M.; Li
 A:Title: Genome sequence of Halobacterium species NRC-1.
 A:Reference number: A84160; MUID:20504483; PMID:11016950
 A:Accession: B84176
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-232 <STO>
 A:Cross-references: GB:AE004437; NID:910579804; PIDN:AA318774.1; GSPDB:GN00138
 C:Genetics:
 A:Gene: VNG0156C

Query Match 43.4%; Score 49; DB 2; Length 232;
 Best Local Similarity 47.4%; Pred. No. 3.8;
 Matches 9; Conservative 4; Mismatches 6; Indels 0; Gaps 0;
 QY 1 AIDHLSLEAMSGNDNDYV 19
 DB 62 AADPVRSVDAMSGRDADHV 80

RESULT 13
 ISBYT2
 DNA topoisomerase (ATP-hydrolyzing) (EC 5.99.1.3) - yeast (Saccharomyces cerevisiae)
 N:Alternate names: DNA gyrase; DNA topoisomerase II; protein N2244; protein YNL088w
 C:Species: Saccharomyces cerevisiae
 C:Date: 30-Sep-1992 #sequence_revision 10-Nov-1995 #text_change 16-Jun-2000
 C:Accession: S57534; A25630; S63027; S30866; S65093
 R:Solier-Mira, A.; Saliz, J.B.; Ballesta, J.P.G.; Remacha, M.
 submitted to the EMBL Data Library, June 1995
 A:Reference number: S57533
 A:Accession: S57534
 A:Molecule type: DNA
 A:Residues: 1-1428 <SOL>
 A:Cross-references: EMBL:X89016; NID:9887621; PIDN:CAA61422.1; PID:9887623
 R:Glaever, G.; Lynn, R.; Goto, T.; Wang, J.C.
 U:BIOL. Chem. 261, 12448-12454, 1986
 A:Title: The complete nucleotide sequence of the structural gene TOP2 of yeast DNA topoi
 A:Reference number: A25630; MUID:86304413; PMID:3017975
 A:Accession: A25630
 A:Molecule type: DNA
 A:Residues: 1-74, 'N', 75-546, 'L', 548-836, 'R', 838-1428 <GIA>
 A:Cross-references: GB:M1814; NID:9172997; PIDN:AA336510.1; PID:9172998
 R:Solier-Mira, A.; Saliz, J.B.; Ballesta, J.P.G.; Remacha, M.
 submitted to the Protein Sequence Database, April 1996
 A:Reference number: S63018
 A:Accession: S63018
 A:Molecule type: DNA

A:Residues: 1-1428 <SOM>
 A:Cross-references: EMBL:Z71364; NID:91301988; PIDN:CAA95964.1; PID:91301989; GSPDB:GN00
 A:Experimental source: strain S288C
 R:Fannatipour, M.; Liu, Y.X.; Nilles, J.L.
 submitted to the EMBL Data Library, January 1993
 A:Description: The top2-5 mutant of yeast topoisomerase II encodes an enzyme resistant t
 A:Reference number: S30866

A:Accession: S30866
 A:Molecule type: DNA
 A:Residues: 812-836, 'R', 838-882, 'P', 884, 'IT', 887-977 <JAN>
 A:Cross-references: EMBL:D08968; NID:9172999; PIDN:AA859328.1; PID:9173000
 R:Solier-Mira, A.; Saliz, J.E.; Ballesta, J.P.G.; Remacha, M.
 Yeast 12, 485-491, 1996
 A:Title: The sequence of a 17 933 bp segment of Saccharomyces cerevisiae chromosome XIV
 A:Reference number: S65092; MUID:196310628; PMID:8740422

A:Accession: S65093
 A:Status: nucleic acid sequence not shown
 A:Molecule type: DNA
 A:Residues: 1-1428 <SOF>
 A:Cross-references: EMBL:X89016; NID:9887621; PIDN:CAA61422.1; PID:9887623
 C:Comment: Type II DNA topoisomerase catalyzes the ATP-dependent transient breakage, pass
 C:Genetics:
 A:Gene: SGD:TOP2; TOR3; TRF3; TOP2-5; MIPS:YNL088w
 A:Map position: 14L
 C:Superfamily: eukaryotic type II DNA topoisomerase; phase T4 DNA topoisomerase (ATP-hyc
 C:Keywords: ATP; DNA binding; DNA replication; heterotetramer; isomerase; phosphoprotein

Query Match 43.4%; Score 49; DB 1; Length 1428;
 Best Local Similarity 50.0%; Pred. No. 30;
 Matches 8; Conservative 2; Mismatches 6; Indels 0; Gaps 0;
 QY 4 HSLSLAMSGNDNDYV 19
 DB 623 HLKIFSLQGNDRDYI 638

RESULT 14
 AE3211
 Hypothetical protein potF [imported] - Agrobacterium tumefaciens (strain C58, Dupont) p1
 C:Species: Agrobacterium tumefaciens
 C:Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 18-Nov-2002
 C:Accession: AE3211
 R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, L
 erage, G.; Gillet, W.; Grant, C.; Gunthner, D.; Kutayavin, T.; Levy, R.; Li, M.; McCell
 Science 294, 2317-2323, 2001
 A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Tung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,
 ster, B.W.
 A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
 A:Reference number: AB2577; MUID:21608550; PMID:11743193
 A:Accession: AE3211
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-366 <KUR>
 A:Cross-references: GB:AE008687; PIDN:AA46107.1; PID:917743872; GSPDB:GN00188
 A:Experimental source: strain C58 (Dupont)
 C:Genetics:
 A:Gene: potF
 A:Genome: plasmid

Query Match 42.5%; Score 48; DB 2; Length 366;
 Best Local Similarity 52.9%; Pred. No. 9.2;
 Matches 9; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 3 DHSLSLEAMSGNDNDYV 19
 DB 179 DFLSLMLYQGNDRDYV 195
 RESULT 15
 T04254
 hypothetical protein F20B18.100 - Arabidopsis thaliana

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 13, 2003, 09:31:40 ; Search time 7.56398 Seconds

(without alignments)
130.561 Million cell updates/sec

Title: US-09-290-049a-13

Perfect score: 113
Sequence: 1 AIDHISILEANSQNDNDYKQ 21

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	108	95.6	1365	1 GTF5_STRDO	P29336 streptococc
2	75	66.4	1462	1 GTFB_STRMU	P49331 streptococc
3	73	64.6	1476	1 GTFB_STRMU	P08887 streptococc
4	71	62.8	1455	1 GTFB_STRMU	P13470 streptococc
5	62	54.9	1592	1 GTF2_STRDO	P27470 streptococc
6	62	54.9	1597	1 GTF1_STRDO	P11001 streptococc
7	49	43.4	1428	1 TOP2_YEAST	P06786 saccharomyc
8	48	42.5	1306	1 POLG_PSBMV	P29152 p genome po
9	48	42.5	6885	1 SNE2_HUMAN	O87400 homo sapien
10	47	41.6	680	1 YH2X_SCHPO	O74443 schizosacch
11	45	39.8	2365	1 TOXB_CLODI	P18177 clostridium
12	44	38.9	341	1 AC0B_BACSV	O34591 bacillus su
13	44	38.9	487	1 STK4_HUMAN	O13043 homo sapien
14	43	38.1	572	1 PKMT_DICDI	O23919 dictyostell
15	43	38.1	674	1 COMA_BACTU	O87905 bacillus th
16	43	38.1	766	1 SC15_DROME	O94546 drosophila
17	43	38.1	1161	1 RPO2_FOMPV	O91544 fowlpox vir
18	42.5	37.6	272	1 Y028_BUCAL	P57140 buchnera ap
19	42	37.2	794	1 Y966_CORGL	P53667 corynebacte
20	41.5	36.7	128	1 Y532_BUCAL	P57598 buchnera ap
21	41.5	36.7	151	1 Y347_METUA	O57793 methanococc
22	41.5	36.7	680	1 NCPK_CANTR	P37201 candida tro
23	41.5	36.7	844	1 PHSG_DROME	O94519 drosophila
24	41	36.3	116	1 Y960_HAEIN	P44084 haemophilus
25	41	36.3	295	1 VENN_VIBVU	P74965 vibrio vln
26	41	36.3	396	1 PKG_TEPIN	O85818 leprospira
27	41	36.3	666	1 POLI_FMYD	P09523 figwort mos
28	41	36.3	801	1 PIR4_DICDI	O02158 dictyostell
29	41	36.3	808	1 PHK1_ANASP	O81481 anabaena sp
30	41	36.3	923	1 K685_MOUSE	O81392 mus musculi
31	41	36.3	927	1 K685_HUMAN	O75170 homo sapien
32	41	36.3	1141	1 UBP2_SCHPO	O93100 schizosacch
33	41	36.3	1634	1 DPOL_METUA	O58255 methanococc

34	40.5	35.8	310	1	PIR1_ARATH	O91ps9 arabidopsis
35	40.5	35.8	833	1	CM41_YEAST	P53008 saccharomyc
36	40.5	35.8	1461	1	TOP2_CANAL	P87078 candida alb
37	40	35.4	53	1	YORU_TTVI	P19305 thermoprote
38	40	35.4	121	1	RUS_SOLME	P93779 solanum mel
39	40	35.4	179	1	FLJA_SALAB	P52618 salmoneella
40	40	35.4	179	1	FLJA_SALAE	P52617 salmoneella
41	40	35.4	179	1	FLJA_SALTY	P52619 salmoneella
42	40	35.4	261	1	UPK_XYLFA	O9pce0 xyella tas
43	40	35.4	263	1	UPK_XANCP	O8pda29 xanthomonas
44	40	35.4	265	1	UPK_XANAC	O8paw6 xanthomonas
45	40	35.4	302	1	CH14_SOLTU	P52406 solanum tub

ALIGNMENTS

RESULT 1	ID	GFES_STRDO	STANDARD	PRT	1365 AA.
AC	P29336				
DT	01-DEC-1992 (Rel. 24, Created)				
DT	01-DEC-1992 (Rel. 24, Last sequence update)				
DT	28-FEB-2003 (Rel. 41, Last annotation update)				
DE	Glucosyltransferase-5 precursor (EC 2.4.1.5) (GNF-S) (Dextranucrase)				
DE	(Sucrose 6-glucosyltransferase).				
GN	GTF5.				
OS	Streptococcus downei (Streptococcus sobrinus).				
OC	Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;				
CC	Streptococcus.				
OX	NCBI_TaxID=1317;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=MFE28;				
RX	MEDLINE=90316665; PubMed=2142479;				
RA	Gilmore K.S., Russell R.R., Ferretti J.J.;				
RT	"Analysis of the Streptococcus downei gtf5 gene, which specifies a				
RT	glucosyltransferase that synthesizes soluble glucans.",				
RL	Infect. Immun. 58:2452-2458(1990).				
CC	- FUNCTION: PRODUCTION OF EXTRACELLULAR GLUCANS, THAT ARE THOUGHT				
CC	TO PLAY A KEY ROLE IN THE DEVELOPMENT OF THE DENTAL PLAQUE BECAUSE				
CC	OF THEIR ABILITY TO ADHERE TO SMOOTH SURFACES AND MEDIANE THE				
CC	AGGREGATION OF BACTERIAL CELLS AND FOOD DEBRIS.				
CC	- CATALYTIC ACTIVITY: Sucrose + {(1,6)-alpha-D-glucosyl} (N) = D-				
CC	fructose + {(1,6)-alpha-D-glucosyl} (N+1).				
CC	- ENZYME REGULATION: GLUCAN SYNTHESIS BY GTF-S IS INDEPENDENT OF				
CC	PRIMER GLUCAN UNLIKE GTF-1.				
CC	- DISEASE: DENTAL CARIES.				
CC	- MISCELLANEOUS: GTF-S SYNTHESIZES WATER-SOLUBLE GLUCANS (ALPHA				
CC	1,6-GLUCOSE).				
CC	- SIMILARITY: TO OTHER GLUCOSYLTRANSFERASES AND SOME TO A GLUCAN-				
CC	BINDING PROTEIN FROM S.MUTANS.				
CC	- SIMILARITY: Contains 10 cell wall binding repeats.				
CC	-----				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration				
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -				
CC	the European Bioinformatics Institute. There are no restrictions on its				
CC	use by non-profit institutions as long as its content is in no way				
CC	modified and this statement is not removed. Usage by and for commercial				
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/				
CC	or send an email to license@sib-sib.ch).				
CC	-----				
CC	EMBL; M30943; AAA26898.1; -				
DR	InterPro; IPR002479; CW_binding.				
DR	InterPro; IPR003318; Glyco_hydro_70.				
DR	Pfam; PF01473; CW_binding_1; 8.				
DR	Pfam; PF02324; Glyco_hydro_70; 1.				
KW	Transferase; Glycosyltransferase; Signal; Repeat; Dental caries.				
FT	SIGNAL	1	36		
FT	CHAIN	37	1365		
FT	REPEAT	157	177		
FT	REPEAT	178	197		
FT	DOMAIN	198	1061		

FT REPEAT 1062 1082 CELL WALL BINDING 3.
FT REPEAT 1083 1102 CELL WALL BINDING 4.
FT REPEAT 1150 1169 CELL WALL BINDING 5.
FT REPEAT 1170 1190 CELL WALL BINDING 6.
FT REPEAT 1225 1243 CELL WALL BINDING 7.
FT REPEAT 1289 1308 CELL WALL BINDING 8.
FT REPEAT 1309 1328 CELL WALL BINDING 9.
FT REPEAT 1331 1352 CELL WALL BINDING 10.
SQ SEQUENCE 1365 AA; 151590 MM; 1672965A2B8C476 CXC64;

Query Match 95.6%; Score 108; DB 1; Length 1365;
Best Local Similarity 100.0%; Pred. No. 6.4e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AIDHSLTEAWSGNDNDYK 20
DB 467 AIDHSLTEAWSGNDNDYK 466

RESULT 2
GTFD_STRMU STANDARD; PRT; 1462 AA.
ID GTFD_STRMU STANDARD; PRT; 1462 AA.
AC PA9331; 069383; 069386; 069389; 069392; 069398;
DT 01-FEB-1996 (Rel. 33, Created)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Glucosyltransferase-S precursor (EC 2.4.1.5) (GTF-S) (Dextranucrase)
DS (sucrose 6-glucosyltransferase).
GN GTFD OR SMU.910.
OS Streptococcus mutans.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1309;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=GS-5;
RA MEDLINE=91100958; PubMed=2148600;
RA Honda O., Kato C., Kuramitsu H.K.;
RT "Nucleotide sequence of the Streptococcus mutans gtfD gene encoding
RT the glucosyltransferase-S enzyme";
RL J. Gen. Microbiol. 136:2099-2105(1990).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=MT4239 / Serotype C, MT4245 / Serotype E, MT4251 / Serotype F,
RC MT4467 / Serotype E, and MT8148 / Serotype C;
RA MEDLINE=98231643; PubMed=9570124;
RA Fujiwara T., Terao Y., Hoshino T., Kawabata S., Sobue S.,
RA Kimura S., Hamada S.;
RT "Molecular analyses of glucosyltransferase genes among strains of
RT Streptococcus mutans";
RL FEMS Microbiol. Lett. 161:331-336(1998).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=UA159 / Serotype C;
RX MEDLINE=22295063; PubMed=12397186;
RA Adic D., Mcshan W.M., McLaughlin R.E., Savic G., Chang J.,
RA Carson M.B., Primeaux C., Tian R., Kenton S., Jia H., Lin S., Qian Y.,
RA Li S., Zhu H., Najjar F., Lai H., White U., Roe B.A., Ferretti J.J.;
RT "Genome sequence of Streptococcus mutans UA159, a cariogenic dental
RT pathogen";
RL Proc. Natl. Acad. Sci. U.S.A. 99:14434-14439(2002).
CC -1- FUNCTION: PRODUCTION OF EXTRACELLULAR GLUCANS, THAT ARE THOUGHT
CC TO PLAY A KEY ROLE IN THE DEVELOPMENT OF THE DENTAL PLAQUE BECAUSE
CC OF THEIR ABILITY TO ADHERE TO SMOOTH SURFACES AND MEDIANE THE
CC AGGREGATION OF BACTERIAL CELLS AND FOOD DEBRIS.
CC -1- CATALYTIC ACTIVITY: Sucrose + [(1,6)-alpha-D-glucosyl] (N) = D-
CC fructose + [(1,6)-alpha-D-glucosyl] (N+1).
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- DISEASE: DENTAL CARIES.
CC -1- MISCELLANEOUS: GTF-I SYNTHESIZES WATER-INSOLUBLE GLUCANS (ALPHA
CC 1,3-LINKED GLUCOSE AND SOME 1,6 LINKAGES), GTF-S SYNTHESIZES
CC WATER-SOLUBLE GLUCANS (ALPHA 1,6-GLUCOSE), GTF-ST SYNTHESIZES BOTH
CC FORMS OF GLUCANS.

CC -1- SIMILARITY: TO OTHER GLUCOSYLTRANSFERASES AND SOME TO A GLUCAN-
CC BINDING PROTEIN FROM S. MUTANS.
CC -1- SIMILARITY: Contains 6 cell wall binding repeats.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M29296; AAA26895.1; -
CC EMBL; D88653; BAA26103.1; -
CC DR EMBL; D88656; BAA26107.1; -
CC DR EMBL; D88659; BAA26111.1; -
CC DR EMBL; D88662; BAA26115.1; -
CC DR EMBL; D89979; BAA26121.1; -
CC DR EMBL; AE014932; AAN58619.1; -
CC DR InterPro; IPR002479; CW binding.
CC DR InterPro; IPR003318; Glyco_hydro_70.
CC DR Pfam; PF02324; Glyco_hydro_70; I
CC Transferrase; Glycosyltransferase; Signal; Repeat; Dental caries;
CC Complete proteome.
CC -----
CC SIGNAL 1 1462 ?
CC CHAIN ? 1462 ?
CC DOMAIN 1232 1423
CC REPEAT 1232 1295
CC REPEAT 1296 1359
CC REPEAT 1360 1423
CC VARIANT 10 10
CC FT 19 19
CC FT 58 58
CC FT 68 68
CC FT 81 81
CC FT 113 113
CC FT 122 122
CC FT 132 132
CC FT 135 135
CC FT 137 137
CC FT 202 202
CC FT 255 255
CC FT 275 275
CC FT 288 288
CC FT 301 301
CC FT 313 313
CC FT 317 317
CC FT 328 328
CC FT 350 350
CC FT 628 633
CC FT 688 688
CC FT 726 732
CC FT 730 730
CC FT 762 762
CC FT 964 964
CC FT 1019 1019
CC FT 1059 1060
CC FT 1060 1060
CC FT 1080 1080
CC FT 1142 1142
CC FT 1198 1198
CC FT 1220 1220
CC FT 1280 1280
CC FT 1282 1282

-1- H (IN STRAINS GS-5, MT4239, MT4245,
MT4251, MT4467 AND MT8148).
I -> V (IN STRAINS GS-5, MT4239, MT4245,
MT4251, MT4467 AND MT8148).
K -> E (IN STRAIN MT4467).
A -> S (IN STRAINS MT4239 AND MT4245).
A -> T (IN STRAINS MT4251 AND MT8148).
T -> I (IN STRAINS MT4239 AND MT4245).
A -> V (IN STRAINS MT4239, MT4245 AND
MT8148).
A -> S (IN STRAINS GS-5 AND MT4467).
A -> V (IN STRAIN MT4245).
A -> T (IN STRAINS GS-5, MT4239, MT4245,
MT4251, MT4467 AND MT8148).
V -> L (IN STRAIN MT4239).
D -> N (IN STRAIN MT8148).
E -> D (IN STRAINS MT4239, MT4245 AND
MT4251).
D -> N (IN STRAINS MT4239, MT4245 AND
MT4251).
Q -> H (IN STRAIN MT4245).
D -> N (IN STRAINS MT4239 AND MT4251).
E -> K (IN STRAIN MT4239).
V -> F (IN STRAIN MT4239).
F -> L (IN STRAINS MT4239, MT4251 AND
MT4467).
KKYATQ -> EKEYTL (IN STRAIN MT4251).
A -> S (IN STRAIN MT4239).
TDGSEA -> ADKNGDS (IN STRAIN MT4251).
TDGGS -> ADKGN (IN STRAINS MT4239 AND
MT4245).
T -> A (IN STRAINS GS-5, MT4239, MT4245,
MT4251, MT4467 AND MT8148).
D -> Y (IN STRAIN MT4251).
E -> K (IN STRAINS MT4245 AND MT4251).
LG -> IR (IN STRAIN MT4251).
G -> R (IN STRAIN MT4245).
G -> Q (IN STRAIN MT4239).
H -> R (IN STRAIN GS-5).
S -> C (IN STRAIN MT4239).
Y -> C (IN STRAINS MT4251 AND MT4467).
F -> L (IN STRAIN MT4467).
Q -> P (IN STRAIN MT4245).


```

FT  VARIANT      1402      1402      S -> G (IN STRAINS GS-5, MT4239, MT4467
FT  VARIANT      1459      1459      AND MT8148).
FT  CONFLICT      570      570      Y -> H (IN STRAIN MT4467).
FT  CONFLICT      800      817      R -> A (IN REF. 1).
FT  CONFLICT      800      817      ADOXVVAASSTPSTGK -> LKXFAFLRLARPHQWA
FT  CONFLICT      1310      1310      (IN REF. 1).
FT  CONFLICT      1476 AA, 165846 MM; 9C6E09F731B4C6CF CRC64;
SQ  SEQUENCE      1476 AA, 1310      H -> L (IN REF. 1)

Query Match      64.6%; Score 73; DB 1; Length 1476;
Best Local Similarity 73.7%; Pred. No. 0.0024;
Matches 14; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY  1 AIDHSLTEAWSGNDNDY 19
Db  481 ANDHSLTEAWSGNDNDY 499

RESULT 4
GTRC_STRMU
ID  GTRC_STRMU      STANDARD; PRT; 1455 AA.
AC  P13470; 069382; 069385; 069388; 069397; P05427;
AT  01-NOV-1988 (Rel. 09, Created)
DT  28-FEB-2003 (Rel. 41, Last sequence update)
DT  15-SEP-2003 (Rel. 42, Last annotation update)
DE  Glucosyltransferase-SI precursor (EC 2.4.1.5) (GTF-SI)
DE  (Dextranucrase) (Sucrose 6-glucosyltransferase).
GN  GTF OR SMU.1005.
OS  Streptococcus mutans.
OC  Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC  Streptococcus.
OX  NCBI_TaxID=1309;
RV  [1]
RV  SEQUENCE FROM N.A.
RP  STRAIN=GS-5;
RX  MEDLINE=89137980; PubMed=2976010;
RA  Ueda S., Shiroza T., Kuramitsu H.K.;
RT  "Sequence analysis of the gtfC gene from Streptococcus mutans GS-5.";
RL  Gene 691101-109(1988).
RN  [2]
RN  SEQUENCE FROM N.A.
RP  STRAIN=MT4239 / Serotype C, MT4245 / Serotype E, MT4251 / Serotype F,
RC  MT4467 / Serotype E, and MT8148 / Serotype C;
RX  MEDLINE=96231643; PubMed=9570124;
RA  Fujiwara T., Terao Y., Hoshino T., Kawabata S., Ooshima T., Sobue S.,
RA  Kimura S., Hamada S.;
RT  "Molecular analyses of glucosyltransferase genes among strains of
RT  Streptococcus mutans.";
RL  FEMS Microbiol. Lett. 161:331-336(1998).
RN  [3]
RN  SEQUENCE FROM N.A.
RP  STRAIN=UAI59 / ATCC 700610 / Serotype C;
RX  MEDLINE=22295063; PubMed=12397186;
RA  Ajdic D., Mcshan W.M., McLaughlin R.E., Savic G., Chang J.,
RA  Carson M.B., Primeaux C., Tian R., Kenton S., Jia H., Lin S., Qian Y.,
RA  Li S., Zhu H., Nejar F., Lai H., White J., Roe B.A., Ferretti J.J.;
RT  "Genome sequence of Streptococcus mutans UAI59, a cariogenic dental
RT  pathogen.";
RL  Proc. Natl. Acad. Sci. U.S.A. 99:14434-14439(2002).
RN  [4]
RN  SEQUENCE OF 1-349 FROM N.A.
RP  STRAIN=GS-5;
RX  MEDLINE=87308013; PubMed=3040685;
RA  Shiroza T., Ueda S., Kuramitsu H.K.;
RT  "Sequence analysis of the gtfB gene from Streptococcus mutans.";
RN  J. Bacteriol. 169:4263-4270(1987).
CC  -1- FUNCTION: PRODUCTION OF EXTRACELLULAR GLUCANS, THAT ARE THOUGHT
CC  TO PLAY A KEY ROLE IN THE DEVELOPMENT OF THE DENTAL PLAQUE BECAUSE
CC  OF THEIR ABILITY TO ADHERE TO SMOOTH SURFACES AND MEDIANE THE
CC  AGGREGATION OF BACTERIAL CELLS AND FOOD DEBRIS.
CC  -1- CATALYTIC ACTIVITY: Sucrose + ((1,6)-alpha-D-glucosyl) (N) = D-
CC  fructose + ((1,6)-alpha-D-glucosyl) (N+1).
CC  -1- SUBCELLULAR LOCATION: Secreted.

```

```

CC  -1- DISEASE: DENTAL CARIES.
CC  -1- MISCELLANEOUS: GTF-1 SYNTHESIZES WATER-INSOLUBLE GLUCANS (ALPHA
CC  1,3-LINKED GLUCOSE AND SOME 1,6 LINKAGES), GTF-S SYNTHESIZES
CC  WATER-SOLUBLE GLUCANS (ALPHA 1,6-GLUCOSE). GTF-SI SYNTHESIZES BOTH
CC  FORMS OF GLUCANS.
CC  -1- SIMILARITY: TO OTHER GLUCOSYLTRANSFERASES AND SOME TO A GLUCAN-
CC  BINDING PROTEIN FROM S. MUTANS.
CC  -1- SIMILARITY: Contains 5 cell wall binding repeats.
CC  -1- SIMILARITY: Contains 5 cell wall binding repeats.
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
DR  EMBL; M22054; AAA88592.1; -
DR  EMBL; D88652; BAA26102.1; -
DR  EMBL; D88655; BAA26106.1; -
DR  EMBL; D88658; BAA26110.1; -
DR  EMBL; D88661; BAA26114.1; -
DR  EMBL; D89978; BAA26120.1; -
DR  EMBL; AE014940; AAN58706.1; -
DR  EMBL; M17361; AAA88589.1; -
DR  InterPro; IPR002479; CW binding.
DR  InterPro; IPR003118; Glyco_hydro_70.
DR  Pfam; PF01473; CW binding 1; 1.
DR  Pfam; PF02324; Glyco_hydro_70; 1.
KW  Transferase; Glucosyltransferase; signal; Repeat; Dental caries;
KW  Complete proteome.
FT  SIGNAL      1      34
FT  CHAIN      35      1455
FT  DOMAIN      35      1050
FT  DOMAIN      1126      1455
FT  DOMAIN      1126      1455
FT  REPEAT      1126      1159
FT  REPEAT      1169      1200
FT  REPEAT      1227      1238
FT  REPEAT      1253      1303
FT  REPEAT      1318      1330
FT  REPEAT      21      21
FT  VARIANT      81      81
FT  VARIANT      106      106
FT  VARIANT      116      116
FT  VARIANT      126      126
FT  VARIANT      150      151
FT  VARIANT      256      256
FT  VARIANT      425      425
FT  VARIANT      519      519
FT  VARIANT      538      538
FT  VARIANT      545      545
FT  VARIANT      597      597
FT  VARIANT      600      600
FT  VARIANT      601      601
FT  VARIANT      614      614
FT  VARIANT      727      727
FT  VARIANT      734      734
FT  VARIANT      964      964
FT  VARIANT      1113      1113
FT  VARIANT      1118      1118
FT  VARIANT      1204      1204
FT  VARIANT      1208      1208
FT  VARIANT      1292      1294
FT  VARIANT      1305      1369
FT  VARIANT      1326      1326
FT  VARIANT      1331      1331
FT  VARIANT      1331      1331

```

```

FT VARIANT 1377 1377 R -> K (IN STRAIN MT8148).
FT VARIANT 1398 1398 V -> I (IN STRAIN MT8148).
FT VARIANT 1424 1424 D -> N (IN STRAIN MT4239).
FT VARIANT 1439 1439 V -> I (IN STRAINS MT4239 AND MT8148).
FT VARIANT 1444 1444 S -> P (IN STRAIN MT8148).
FT CONFLICT 1337 1455 ORLYKSNQAKGKHLITERGRKLYDPNSGNEVRYR
TSSGMYTFGNDGALIGMHVHGKRVYFDENGVRYASHD
QRNMDYDRDFGRGSSAVRFRSRNGFENPERF ->
HASILSLMVFRLRSSLSQSVKVSNTMLLPEMKFVLM
(IN REF. 1).
SQ SEQUENCE 1455 AA; 162965 MW; 3CB455A99A4FEC86 CRC64;

Query Match 62.8%; Score 71; DB 1; Length 1455;
Best Local Similarity 73.7%; Pred. No. 0.0048;
Matches 14; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

OY 1 AIDHSLTEAWSGNDNDYV 19
Db 507 ANDHSLTEAWSGNDNDYV 525

RESULT 5
GTF2_STRDO STANDARD; PRT; 1592 AA.
ID GTF1_STRDO
AC P27470;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Glucosyltransferase-I precursor (EC 2.4.1.5) (GTF-I) (Dextranucrase)
DE (Sucrose 6-glucosyltransferase)
OS Streptococcus downei (Streptococcus sobrinus).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1317;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=6715 / Serotype G;
RX MEDLINE=91123227; PubMed=1704006;
RA Abo H., Matsumura T., Kodama T., Ohta H., Kato K.,
RA Kagawa H.;
RT "Peptide sequences for sucrose splitting and glucan binding within
RT Streptococcus sobrinus glucosyltransferase (water-insoluble glucan
RT synthetase).";
RT J. Bacteriol. 173:989-996(1991).
RL J. Bacteriol. 173:989-996(1991).
CC -1- FUNCTION: PRODUCTION OF EXTRACELLULAR GLUCANS, THAT ARE THOUGHT
CC TO PLAY A KEY ROLE IN THE DEVELOPMENT OF THE DENTAL PLAQUE BECAUSE
CC OF THEIR ABILITY TO ADHERE TO SMOOTH SURFACES AND MEDIATE THE
CC AGGREGATION OF BACTERIAL CELLS AND FOOD DEBRIS.
CC CATALYTIC ACTIVITY: Sucrose + [(1,6)-alpha-D-glucosyl] (N) = D-
CC fructose + [(1,6)-alpha-D-glucosyl] (N+1).
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- DISEASE: DENTAL CARIES.
CC -1- MISCELLANEOUS: GTF-I SYNTHESIZES WATER-INSOLUBLE GLUCANS (ALPHA
CC 1,3-LINKED GLUCOSE AND SOME 1,6 LINKAGES), GTF-S SYNTHESIZES
CC WATER-SOLUBLE GLUCANS (ALPHA 1,6-GLUCOSE). GTF-SI SYNTHESIZES BOTH
CC FORMS OF GLUCANS.
CC -1- SIMILARITY: TO OTHER GLUCOSYLTRANSFERASES AND SOME TO A GLUCAN-
CC BINDING PROTEIN FROM S.MUTANS.
CC -1- SIMILARITY: Contains 16 cell wall binding repeats.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; D90213; BAA14241.1;
CC InterPro; IPR002479; CW binding.
CC InterPro; IPR003318; Glyco_Hydro_70.
CC Pfam; PF02324; Glyco_Hydro_70; 1.

```

```

KW Transferase; Glycosyltransferase; Signal; Repeat; Dental caries.
FT SIGNAL 1 38 POTENTIAL.
FT CHAIN 39 1592 GLUCOSYLTRANSFERASE-I.
FT DOMAIN 39 1044 CATALYTIC (APPROXIMATE).
FT DOMAIN 1093 1592 GLUCAN-BINDING (APPROXIMATE).
FT DOMAIN 1093 1592 6.5 X TANDEM REPEATS.
FT REPEAT 1093 1142 1.
FT REPEAT 1158 1207 2.
FT REPEAT 1222 1272 3.
FT REPEAT 1287 1337 4.
FT REPEAT 1402 1451 5.
FT REPEAT 1514 1563 6.
FT REPEAT 1577 1592 7 (INCOMPLETE).
SQ SEQUENCE 1592 AA; 176167 MW; BCOA6D079351ECF CRC64;

Query Match 54.9%; Score 62; DB 1; Length 1592;
Best Local Similarity 57.9%; Pred. No. 0.14;
Matches 11; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

OY 1 AIDHSLTEAWSGNDNDYV 19
Db 477 ANHSLTEAWSGNDNDYV 495

RESULT 6
GTF1_STRDO STANDARD; PRT; 1597 AA.
ID GTF2_STRDO
AC P11001;
DT 01-JUL-1989 (Rel. 11, Created)
DT 01-JUL-1989 (Rel. 11, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Glucosyltransferase-I precursor (EC 2.4.1.5) (GTF-I) (Dextranucrase)
DE (Sucrose 6-glucosyltransferase).
GN GTFI.
OS Streptococcus downei (Streptococcus sobrinus).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1317;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MFE28;
RX MEDLINE=87308014; PubMed=3040686;
RA Ferretti J.J., Gilpin M.L., Russell R.R.B.;
RT "Nucleotide sequence of a glucosyltransferase gene from Streptococcus
RT sobrinus MFE28.";
RL J. Bacteriol. 169:4271-4278(1987).
CC -1- FUNCTION: PRODUCTION OF EXTRACELLULAR GLUCANS, THAT ARE THOUGHT
CC TO PLAY A KEY ROLE IN THE DEVELOPMENT OF THE DENTAL PLAQUE BECAUSE
CC OF THEIR ABILITY TO ADHERE TO SMOOTH SURFACES AND MEDIATE THE
CC AGGREGATION OF BACTERIAL CELLS AND FOOD DEBRIS.
CC CATALYTIC ACTIVITY: Sucrose + [(1,6)-alpha-D-glucosyl] (N) = D-
CC fructose + [(1,6)-alpha-D-glucosyl] (N+1).
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- DISEASE: DENTAL CARIES.
CC -1- MISCELLANEOUS: GTF-I SYNTHESIZES WATER-INSOLUBLE GLUCANS (ALPHA
CC 1,3-LINKED GLUCOSE AND SOME 1,6 LINKAGES), GTF-S SYNTHESIZES
CC WATER-SOLUBLE GLUCANS (ALPHA 1,6-GLUCOSE). GTF-SI SYNTHESIZES BOTH
CC FORMS OF GLUCANS.
CC -1- SIMILARITY: TO OTHER GLUCOSYLTRANSFERASES AND SOME TO A GLUCAN-
CC BINDING PROTEIN FROM S.MUTANS.
CC -1- SIMILARITY: Contains 19 cell wall binding repeats.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M7391; AAC63063.1;
CC InterPro; IPR002479; CW binding.
CC InterPro; IPR003318; Glyco_Hydro_70.

```

```

DR PFam: PF01473; CW binding_1.16.
DR PFam: PF02324; Glyco_hydro_70; 1.
KW Transferrase; Glycosyltransferase; Signal; Repeat; Dental caries.
FT SIGNAL 1 38
FT CHAIN 39 1597 GLUCOSYLTRANSFERASE-1
FT DOMAIN 39 1050 POTENTIAL.
FT DOMAIN 1099 1597 CATALYTIC (APPROXIMATE).
FT REPEAT 1099 1597 GLUCAN-BINDING (APPROXIMATE).
FT REPEAT 1099 1132 1.25 A, 2 B, AND 5 AC REPEATS.
FT REPEAT 1163 1213 A REPEAT.
FT REPEAT 1227 1277 AC REPEAT.
FT REPEAT 1282 1342 AC REPEAT.
FT REPEAT 1352 1399 AC REPEAT.
FT REPEAT 1406 1455 B REPEAT.
FT REPEAT 1465 1512 AC REPEAT.
FT REPEAT 1519 1568 B REPEAT.
FT REPEAT 1582 1597 AC REPEAT.
SQ SEQUENCE 1597 AA; 177080 MW; B9B6A200868798E CRC64;

Query Match 54.9%; Score 62; DB 1; Length 1597;
Best Local Similarity 57.9%; Pred. No. 0.14;
Matches 11; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 1 AIDHSLIEFAMSGNDIV 19
Db 483 ANNHVSIVEAMSDNDTPYL 501

RESULT 7
TOP2_YEAST STANDARD; PRT; 1428 AA.
AC P06786;
DT 01-JAN-1988 (Rel. 06, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE DNA topoisomerase II (EC 5.99.1.3).
TOP2 OR TOR3 OR YNL088W OR N2244.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=86304413; PubMed=3017975;
RA Glaeser F., Lynn R., Goto T., Wang J.C.;
RT "The complete nucleotide sequence of the structural gene TOP2 of
RT yeast DNA topoisomerase II."
RL J. Biol. Chem. 261:12448-12454(1986).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=S288c / FY1679;
RX MEDLINE=96310628; PubMed=8740422;
RA Soler-Mira A., Saitz J.E., Ballestra J.P.G., Remacha M.;
RT "The sequence of a 17,933 bp segment of Saccharomyces cerevisiae
RT chromosome XIV contains the RHO2, TOP2, MKT1 and ENDS genes and five
RT new open reading frames."
RL Yeast 12:485-491(1996).
RN [3]
RP REVIEW ON PHOSPHORYLATION.
RX MEDLINE=93073815; PubMed=1332607;
RX Glaeser S.M., Walther R., Dang O., Cardenas M.E.;
RA Berger J.M., Gamblin S.J., Harrison S.C., Wang J.C.;
RT "Structure and mechanism of DNA topoisomerase II."
RL Nature 379:225-232(1996).
RN [4]
RP X-RAY CRYSTALLOGRAPHY (2.7 ANGSTROMS) OF 410-1202.
RX MEDLINE=96138378; PubMed=8538787;
RA Berger J.M., Gamblin S.J., Harrison S.C., Wang J.C.;
RT "Structure and mechanism of DNA topoisomerase II."
RL Nature 379:225-232(1996).
CC -1- FUNCTION: CONTROL OF TOPOLOGICAL STATES OF DNA BY TRANSIENT
CC BREAKAGE AND SUBSEQUENT REJOINING OF DNA STRANDS. TOPOISOMERASE II
CC MAKES DOUBLE-STRAND BREAKS. ESSENTIAL DURING MITOSIS AND MEIOSIS
CC FOR PROPER SEGREGATION OF DAUGHTER CHROMOSOMES.

```

FT	HELIX	450	464
FT	STRAND	469	474
FT	HELIX	489	502
FT	TURN	503	505
FT	TURN	509	510
FT	TURN	513	514
FT	STRAND	520	524
FT	HELIX	533	545
FT	HELIX	547	550
FT	TURN	551	551
FT	TURN	553	554
FT	STRAND	556	558
FT	STRAND	564	569
FT	STRAND	573	578
FT	HELIX	581	590
FT	TURN	591	591
FT	HELIX	592	594
FT	STRAND	597	601
FT	HELIX	610	627
FT	TURN	628	629
FT	STRAND	633	691
FT	STRAND	694	694
FT	TURN	695	697
FT	HELIX	701	712
FT	TURN	713	714
FT	STRAND	719	720
FT	HELIX	721	732
FT	TURN	733	733
FT	HELIX	738	749
FT	TURN	753	754
FT	TURN	760	761
FT	STRAND	764	764
FT	HELIX	770	772
FT	TURN	773	773
FT	TURN	775	776
FT	TURN	780	782
FT	STRAND	784	785
FT	TURN	789	790
FT	TURN	791	794
FT	HELIX	797	802
FT	STRAND	805	805
FT	TURN	809	810
FT	STRAND	815	825
FT	HELIX	824	828
FT	STRAND	830	833
FT	STRAND	838	841
FT	STRAND	844	844
FT	HELIX	846	858
FT	TURN	859	859
FT	STRAND	867	867
FT	TURN	870	871
FT	STRAND	875	878
FT	TURN	881	882
FT	STRAND	883	887
FT	STRAND	889	894
FT	TURN	895	896
FT	STRAND	897	902
FT	TURN	905	906
FT	HELIX	909	920
FT	STRAND	932	935
FT	STRAND	943	946
FT	HELIX	949	958
FT	HELIX	960	963
FT	TURN	964	965
FT	STRAND	967	971
FT	TURN	973	974
FT	STRAND	975	978
FT	TURN	980	981
FT	STRAND	984	987
FT	HELIX	990	1034
FT	TURN	1035	1036
FT	HELIX	1045	1054
FT	TURN	1055	1056

FT STRAND 1059 1060
FT TURN 1062 1063
FT STRAND 1066 1067
FT TURN 1108 1110
FT STRAND 1112 1112
FT HELIX 1114 1117
FT TURN 1116 1118
FT HELIX 1121 1124
FT HELIX 1126 1147
FT TURN 1148 1148
FT HELIX 1151 1175

Query Match 43.4%; Score 49; DB 1; Length 1428;
Best Local Similarity 50.0%; Pred. No. 14;
Matches 8; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

4 HSLTEAMSGNDNDYV 19
623 HLKIPHSLOGNDKDYI 638

Db

RESULT 8
POLG_PSBMV STANDARD; PRT; 3206 AA.
AC P29152;
ID POLG_PSBMV
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Genome polypeptide [contains: N-terminal protein (p1); Helper
component proteinase (EC 3.4.22.45) (HC-Pro); Protein p3; 6 kDa
protein 1 (6k1); Cytoplasmic inclusion protein (CI); 6 kDa protein 2
(6K2); Genome-linked protein (VPG); Nuclear inclusion protein A (NI-A)
(NIA) (EC 3.4.22.44) (49 kDa proteinase) (49 kDa-Pro); Nuclear
inclusion protein B (NI-B) (NIB) (RNA-directed RNA polymerase)
(EC 2.7.7.48); Coat protein (CP)].
DE Pea seed-borne mosaic virus (strain DPDI).
OS Viruses; ssRNA positive-strand viruses, no DNA stage; Polyviridae;
OC Polyvirus.
OC NCBI_TaxID=31736;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92044431; PubMed=1940858;
RA Johansen E., Rasmussen O.F., Heide M., Borkhardt B.;
RT "The complete nucleotide sequence of pea seed-borne mosaic virus
RNA.";
RL J. Gen. Virol. 72:2625-2632 (1991).
CC -1- FUNCTION: HELPER COMPONENT-PROTEINASE IS REQUIRED FOR APHID
TRANSMISSION AND ALSO HAS PROTEOLYTIC ACTIVITY.
CC -1- FUNCTION: CYTOPLASMIC INCLUSION PROTEIN HAS HELICASE ACTIVITY. IT
MAY BE INVOLVED IN REPLICATION.
CC -1- FUNCTION: NUCLEAR INCLUSION PROTEIN A HAS PROTEOLYTIC ACTIVITY.
CC -1- CATALYTIC ACTIVITY: Hydrolyzes glutamyl bonds, and activity is
further restricted by preferences for the amino acids in p6 - p1,
that vary with the species of polyvirus, e.g. Glu-Xaa-Xaa-Tyr-Xaa-
Glu+(Ser or Gly) for the enzyme from tobacco etch virus. The
natural substrate is the viral polypeptide, but other proteins and
oligopeptides containing the appropriate consensus sequence are
also cleaved.
CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
{RNA} (N).
CC -1- CATALYTIC ACTIVITY: Hydrolyzes a Gly-|Gly bond at its own C-
terminus, commonly in the sequence -Tyr-Xaa-Val-Gly-|Gly, in the
processing of the polyviral polyprotein.
CC -1- PTM: VPG IS COVALENTLY LINKED TO THE GENOMIC RNA.
CC -1- PTM: THE VIRAL RNA OF POTYVIRUSES IS EXPRESSED AS A SINGLE
POLYPEPTIDE WHICH UNDERGOES POSTTRANSLATIONAL PROTEOLYTIC
PROCESSING RESULTING IN THE PRODUCTION OF AT LEAST EIGHT
INDIVIDUAL PROTEINS.
CC -1- SIMILARITY: HC PROTEINASE BELONGS TO PEPTIDASE FAMILY C6.
CC -1- SIMILARITY: NI-A PROTEINASE BELONGS TO PEPTIDASE FAMILY C4.
CC -1- SIMILARITY: BELONGS TO THE POTYVIRUSES POLYPEPTIDE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@eb-sib.ch).

DR EMBL: D10930; BAA01726.1; -
 DR PIR: J01331; GNVSFV.
 DR MEROPS; C04.010; -
 DR MEROPS; C06.001; -
 DR InterPro; IPR001410; DEAD.
 DR InterPro; IPR001650; Helicase_C.
 DR InterPro; IPR001730; Peptidase_C4.
 DR InterPro; IPR001456; Peptidase_C6.
 DR InterPro; IPR001592; Poly.coat.
 DR InterPro; IPR002540; Poly.pl.
 DR InterPro; IPR007095; RNA.pol.DS.PS.
 DR InterPro; IPR001205; RNA.pol.p3.
 DR InterPro; IPR007094; RNA.pol.psvir.
 DR Pfam; PF00271; helicase_C; 1.
 DR Pfam; PF00863; Peptidase_C4; 1.
 DR Pfam; PF00851; Peptidase_C6; 1.
 DR Pfam; PF00767; Poly.coat; 1.
 DR Pfam; PF01577; Poly.pl; 1.
 DR Pfam; PF00680; RNA_dep.RNA.pol; 1.
 DR PRINTS; PR00986; NIAPO1PTASE.
 DR SMART; SM00487; DEXDC; 1.
 DR SMART; SM00490; HELIC; 1.
 DR Hydroxylase; Transferase; 1.
 KM Coat protein; Polyprotein; Covalent protein-RNA linkage; Helicase;
 KM ATP-binding.
 FT CHAIN 1 ? N-TERMINAL PROTEIN.
 FT CHAIN ? 856 HELPER COMPONENT PROTEINASE.
 FT CHAIN ? ? PROTEIN P3.
 FT CHAIN ? 1266 6 kDa PROTEIN 1.
 FT CHAIN 1267 1902 CYTOPLASMIC INCLUSION PROTEIN.
 FT CHAIN 1903 1955 6 kDa PROTEIN 2.
 FT CHAIN 1956 ? GENOME-LINKED PROTEIN.
 FT CHAIN ? 2395 NUCLEAR INCLUSION PROTEIN A.
 FT CHAIN 2396 2915 NUCLEAR INCLUSION PROTEIN B.
 FT CHAIN 2916 3206 COAT PROTEIN.
 FT CHAIN 1267 1267 CLEAVAGE (BY 49 kDa PROTEASE).
 FT SITE 1266 1267 CLEAVAGE (BY 49 kDa PROTEASE).
 FT SITE 1902 1903 CLEAVAGE (BY 49 kDa PROTEASE).
 FT SITE 1955 1956 CLEAVAGE (BY 49 kDa PROTEASE).
 FT SITE 2395 2396 CLEAVAGE (BY 49 kDa PROTEASE).
 FT SITE 2915 2916 CLEAVAGE (BY 49 kDa PROTEASE).
 FT BINDING 2016 COVALENT LINKAGE OF VIRAL RNA (BY
 FT SIMILARITY).
 SO NP BIND 1351 1358 ATP (POTENTIAL).
 SQ SEQUENCE 3206 AA; 364271 MW; 42A3D921BB940CBF CRC64;
 Query Match 42.5%; Score 48; DB 1; Length 3206;
 Best Local Similarity 47.6%; Pred. No. 51;
 Matches 10; Conservative 4; Mismatches 5; Indels 2; Gaps 1;
 QY 2 IDHLSI--LEAMSGNDNDVYK 20
 DB 394 IDHFSIVGIKINWAFDAEYK 414
 RESULT 9
 SNE2_HUMAN STANDARD; PRT; 6885 AA.
 AC Q8WQH0; Q8WIS3; Q8NF49; Q8TER7; Q8WWM3; Q8WWM4; Q8WWM5; Q8WXM1;
 AC Q9N050; Q9NFC4; Q9Y2L4; Q9YAR1;
 DT 15-SEP-2003 (Rel. 42, Created)
 DT 15-SEP-2003 (Rel. 42, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Nesprin 2 (Nuclear envelope spectrin repeat protein 2) (Syn-2)
 DE (Synaptic nuclear envelope protein 2) (Nucleus and actin connecting
 DE element protein) (NUANCE protein).
 GN SNE2 OR NDA OR KIAA1011.

OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OC NCBI_TaxID:9606;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 9), FUNCTION, CHARACTERIZATION, AND
 RP INTERACTION WITH F-ACTIN.
 RX MEDLINE=22113122; PubMed=12118075;
 RA Zhen Y.-Y., Libotte T., Munck M., Noegel A.A., Korenbaum E.;
 RT "NUANCE, a giant protein connecting the nucleus and actin
 RT cytoskeleton." J. Cell Sci. 115:3207-3222(2002).
 RL J. Cell Sci. 115:3207-3222(2002).
 RN [2]
 RP SEQUENCE FROM N.A. (ISOFORMS 4; 5 AND 7).
 RX MEDLINE=21652858; PubMed=11792814;
 RA Zhang Q., Skepper J.N., Yang F., Davies J.D., Hegyi L., Roberts R.G.,
 RT Weisberg P.L., Ellis J.A., Shanahan C.M.;
 RT "Nesprins: a novel family of spectrin-repeat-containing proteins that
 RT localize to the nuclear membrane in multiple tissues." J.
 RT J. Cell Sci. 114:4485-4498(2001).
 RN [3]
 RP SEQUENCE FROM N.A. (ISOFORM 1).
 RX MEDLINE=22296983; PubMed=12408964;
 RA Zhang Q., Kagnauth C., Greener M.J., Shanahan C.M., Roberts R.G.;
 RT "The nesprins are giant actin-binding proteins, orthologous to
 RT Drosophila melanogaster muscle protein MSP-300." J.
 RT Genomics 80:473-481(2002).
 RN [4]
 RP SEQUENCE FROM N.A. (ISOFORM 6).
 RX TISSUE-Testis.
 RA Pousetka A., Klein M., Mewes H.-W., Gassenhuber J., Wiemann S.;
 RT Submitted (SEP-1999) to the EMBL/Genbank/DBJ databases.
 RL [5]
 RP SEQUENCE FROM N.A.
 RX PubMed=12508121;
 RA Heilig R., Eckenberg R., Petit J.-L., Fonknechten N., Da Silva C.,
 RA Catolico L., Levy M., Barbe V., de Bernardis V., Ureca-vial A.,
 RA Pelletier E., Vico V., Anthouard V., Rowen L., Madan A., Qin S.,
 RA Sun H., Du H., Pepin K., Artiguenave F., Robert C., Cruaud C.,
 RA Brulets T., Jallion O., Friedlander L., Samson G., Broctier P.,
 RA Cure S., Segurens B., Aniere F., Samain S., Crespeau H., Abbasi N.,
 RA Alich N., Bocus D., Dickhoff R., Dors M., Dubois I., Friedman C.,
 RA Gouyenoux M., James R., Madan A., Mairey-Bettrada B., Mangenot S.,
 RA Martins N., Menard M., Oztas S., Ratcliffe A., Shafer T., Traak B.,
 RA Vacherie B., Bellemere C., Belser C., Besnard-Gonnert M.,
 RA Barot-Maveil D., Boucard M., Briex-Silla S., Combette S.,
 RA Dufosse-Laurent V., Ferron C., Lechaplats C., Louesse C., Mousellet D.,
 RA Magdeleat G., Patteu F., Petit E., Stryain-Trukiewicz P., Trybou A.,
 RA Vega-Czarny N., Bataille E., Bluet E., Bordelais I., Dubois M.,
 RA Dumont C., Guerin T., Haflray S., Hammadi R., Muanga J., Pellouin V.,
 RA Robert D., Verdier E., Gauguier G., Roy A., Sainte-Marthe L.,
 RA Verdier J., Verdier-Discala C., Hillier L., Fulton L., Moperson J.,
 RA Matsuda F., Wilson R., Scarpetti C., Gyapay G., Winkler P., Saurin W.,
 RA Querier F., Waterston R., Hood L., Weissbach J.;
 RT "The DNA sequence and analysis of human chromosome 14." J.
 RT Nature 421:601-607(2003).
 RL [6]
 RP SEQUENCE FROM N.A. (ISOFORM 8).
 RX TISSUE=Brain;
 MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altshul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh P.,
 RA Diatchenko L., Marzina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Ueda T.B., Toshitsuki S., Carrini P., Prange C.J.,
 RA Raha S.S., Loggellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.O., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Huiyik S.W.,
 RA Villalón D.K., Muzny K.C., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whitting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,

RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield A.S.N., Krzywinski M.I., Skalska U., Smallus D.E.,
 RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RP [7]
 RP SEQUENCE FROM N.A. (ISOFORM 3), AND SEQUENCE OF 1-956 AND 5133-6885
 RP FROM N.A.
 RC TISSUE=Splice, and Tongue;
 RA Jikuya H., Takano T., Nomura N., Kikuno R., Nagase T., Ohara O.,
 RA Nimomiyu K., Magatsuma M., Kanda K., Kondo H., Yokoi T., Kodaira H.,
 RA Furuya T., Takahashi M., Kikawa E., Omura Y., Abe K., Kamihara K.,
 RA Kasuya N., Sato K., Tanikawa M., Yamazaki M., Sugiyama T., Irie R.,
 RA Otsuki T., Sato H., Makimatsu A., Ishii S., Yamamoto Y., Isono Y.,
 RA Okawaki H., Sato K., Nishikawa T., Kimura K., Yamashita H.,
 RA Matsuo K., Nakamura Y., Sekine M., Kikuchi H., Murakawa K.,
 RA Kanehori K., Takahashi-Fujii A., Oshima A., Sugiyama A., Kawakami B.,
 RA Suzuki Y., Sugano S., Nagahari K., Masuno Y., Nagai K., Isogai T.;
 RT "NDO human cDNA sequencing project";
 RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
 RP [8]
 RP SEQUENCE OF 5709-6885 FROM N.A. (ISOFORM 2).
 RP TISSUE=Brain;
 RC MEDLINE=99246063; PubMed=10231033;
 RA Nagase T., Ishikawa K.-I., Suyama M., Kikuno R., Hirose A.,
 RA Miyajima N., Tanaka A., Kotani H., Nomura N., Ohara O.;
 RT "Prediction of the coding sequences of unidentified human gene. XIII.
 RT The complete sequences of 100 new cDNA clones from brain which code
 RT for large proteins in vitro";
 RL DNA Res. 6:63-70(1999).
 RP [9]
 RP REVISIONS.
 RC TISSUE=Brain;
 RC MEDLINE=22158633; PubMed=12168954;
 RX Nakajima D., Okazaki N., Yamakawa H., Kikuno R., Ohara O., Nagase T.;
 RT "Construction of expression-ready cDNA clones for KIAA genes: manual
 RT curation of 330 KIAA cDNA clones.";
 RL DNA Res. 9:99-106(2002).
 RP [10]
 RP SEQUENCE OF 5754-6885 FROM N.A.
 RP MEDLINE=21154917; PubMed=11250166;
 RX Wiemann S., Weil B., Wellenreuther R., Gaassenhuber J., Glaesl S.,
 RA Ansoyge W., Boecker M., Bloecher H., Bauersachs S., Blum H.,
 RA Lauber U., Duesterhoeft A., Beyer A., Koehrer K., Strack N.,
 RA Mewes H.-W., Oltenevaider B., Obermaier B., Tampe J., Heubner D.,
 RA Mambert R., Korn B., Klein M., Poustka A.;
 RT "Towards a catalog of human genes and proteins: sequencing and
 RT analysis of 500 novel complete protein coding human cDNAs";
 RL Genome Res. 11:422-435(2001).
 CC -1- FUNCTION: Involved in the maintenance of nuclear organization and
 CC structural integrity. Probable anchoring protein which tethers the
 CC nucleus to the cytoskeleton. Connects nuclei to the cytoskeleton
 CC by interacting with the nuclear envelope and with F-actin in the
 CC cytoplasm.
 CC -1- SUBUNIT: Interacts with F-actin via its N-terminal domain.
 CC -1- SUBCELLULAR LOCATION: Type IV membrane protein (Potential). The
 CC largest part of the protein is cytoplasmic, while its C-terminal
 CC part is associated with the nuclear envelope, most probably the
 CC outer nuclear membrane. Remains associated with the nuclear
 CC envelope during its breakdown in mitotic cells.
 CC -1- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named Isoforms=9;
 CC Name=1;
 CC IsoId=Q8WXH0-1; Sequence=Displayed;
 CC Name=2;
 CC IsoId=Q8WXH0-2; Sequence=VSP_007164, VSP_007166;
 CC Note=No experimental confirmation available;
 CC Name=3;
 CC IsoId=Q8WXH0-3; Sequence=VSP_007155;
 CC Note=Produced by exon skipping that results in a frameshift. No
 CC experimental confirmation available;

CC Name=4; Synonyms=Beta;
 CC IsoId=Q8WXH0-4; Sequence=VSP_007156;
 CC Name=5; Synonyms=Alpha;
 CC IsoId=Q8WXH0-5; Sequence=VSP_007157, VSP_007164, VSP_007165;
 CC Name=6;
 CC IsoId=Q8WXH0-6; Sequence=VSP_007158, VSP_007165, VSP_007166;
 CC Note=No experimental confirmation available;
 CC Name=7; Synonyms=Gamma;
 CC IsoId=Q8WXH0-7; Sequence=VSP_007154, VSP_007163;
 CC Name=8;
 CC IsoId=Q8WXH0-8; Sequence=VSP_007161, VSP_007162;
 CC Note=No experimental confirmation available;
 CC Name=9; Synonyms=NUANCE-N-33;
 CC IsoId=Q8WXH0-9; Sequence=VSP_007159, VSP_007160;
 CC -1- TISSUE SPECIFICITY: Widely expressed, with higher level in kidney,
 CC adult and fetal liver, stomach and placenta. Weakly expressed in
 CC skeletal muscle and brain. Isoform 5 is highly expressed in
 CC pancreas, skeletal muscle and heart.
 CC -1- DOMAIN: The Klarsicht domain mediates the nuclear envelope
 CC targeting.
 CC -1- SIMILARITY: Belongs to the Nesprin family.
 CC -1- SIMILARITY: Contains 1 actin-binding domain.
 CC -1- SIMILARITY: Contains 2 calponin-homology (CH) domains.
 CC -1- SIMILARITY: Contains 1 Klarsicht domain.
 CC -1- SIMILARITY: Contains 13 leucine-rich (LRR) repeats.
 CC -1- SIMILARITY: Contains 9 spectrin repeats.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC DR EMBL: AF435010; AAL33547.1; -
 CC DR EMBL: AF435011; AAL33548.1; -
 CC DR EMBL: AY061757; AAL33800.1; -
 CC DR EMBL: AY061758; AAL33801.1; -
 CC DR EMBL: AY061759; AAL33802.1; -
 CC DR EMBL: AF495911; AAN60443.1; -
 CC DR EMBL: AL117404; CAB55905.1; -
 CC DR EMBL: AL117404; CAB55905.1; -
 CC DR EMBL: AL152832; -; NOT_ANNOTATED_CDS.
 CC DR EMBL: AL355094; -; NOT_ANNOTATED_CDS.
 CC -----
 CC Query Match 42.5%; Score 48; DB 1; Length 6885;
 CC Best Local Similarity 57.1%; Pred. No. 1.2e+02;
 CC Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
 CC -----
 CC QY 3 DHTSLHAWSGNDN 16
 CC Db 362 DHTQLREAWDGDH 375
 CC |||:|||||:
 CC YHXX SCHPO STANDARD; PRT; 680 AA.
 CC AC YHXX SCHPO
 CC DT 16-OCT-2001 (Rel. 40, Created)
 CC DT 16-OCT-2001 (Rel. 40, Last sequence update)
 CC DT 15-SEP-2003 (Rel. 42, Last annotation update)
 CC DE Hypothetical UPF0024 protein CIA4.09 in chromosome II.
 CC GN SPBC1A4.09.
 CC OS Schizosaccharomyces pombe (fission yeast).
 CC OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 CC OC Schizosaccharomycetales; Schizosaccharomycetaceae;
 CC OC Schizosaccharomycetes.
 CC OX NCBI_TaxID=4896;
 CC RN [1]
 CC RP SEQUENCE FROM N.A.
 CC RC STRAIN=972;
 CC RX MEDLINE=21848401; PubMed=11859360;
 CC Wood V., Gwilliam R., Kajandream M.A., Lyne M., Lyne R., Stewart A.,

RA Sgouras J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
 RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
 RA Collins M., Connor R., Cronin A., Davis P., Felwell T., Fraser A.,
 RA Gantley S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
 RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagals K.,
 RA James K., Jones L., Jones M., Leather S., McDonald S., McLean U.,
 RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odeh C.,
 RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitch E.,
 RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
 RA Skelton O., Simmonds M., Squares R., Squares S., Stevens K.,
 RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
 RA Woodward J., Volkart G., Aert R., Robben J., Gymnopoulos B.,
 RA Welfens I., Vanstreets E., Rieger M., Schaefer M., Mueller-Auer S.,
 RA Gabel C., Fuchs M., Fritze C., Holzer E., Moestl D., Hilbert H.,
 RA Borzym K., Langer I., Beck A., Lebrach H., Reinhardt R., Pohl T.M.,
 RA Eger P., Zimmermann W., Wedler H., Wambut R., Purnelle B.,
 RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelarge V., Moutier S.,
 RA Gilbert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
 RA Lucas M., Rochet M., Gallardin C., Tallada V.A., Garzon A., Rhode G.,
 RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
 RA Dominguez A., Revuelta J.L., Moreno S., Armstrong U., Forsburg S.L.,
 RA Cerutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
 RA Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.,
 RA "The genome sequence of Schizosaccharomyces pombe."
 RA Nature 415:871-880(2002).
 RL Nature 415:871-880(2002).
 CC -1- SIMILARITY: Belongs to the UPF0024 family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AL031174; CAA20114.1; -
 DR PIR; T39858; T39858.
 DR GeneDB_SPombe; SPBCL4.09; -
 DR InterPro; IPR001656; UPF0024.
 DR Pfam; PF01142; UPF0024; 1.
 DR TIGRPFAMs; TIGR00094; TIGR00094; 1.
 KW Hypothetical protein.
 SQ SEQUENCE 680 AA; 76438 MW; B4A52DAECBA1B9A6 CRC64;
 QY
 Query Match 41.6%; Score 47; DB 1; Length 680;
 Best Local Similarity 62.5%; Pred. No. 13;
 Matches 10; Conservative 1; Mismatches 3; Indels 2; Gaps 1;
 Db 434 SILETWSRSGNCTDYL 449
 6 SILEAW--SGNDNDYV 19
 RESULT 11
 TOXB_CLODI
 ID TOXB_CLODI STANDARD; PRT; 2366 AA.
 AC P18177;
 DT 01-NOV-1990 (Rel. 16, Created)
 DT 01-NOV-1990 (Rel. 16, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Toxin B.
 GN TOXB OR TCDB.
 OS Clostridium difficile.
 OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
 OC Clostridium.
 OX NCBI_TaxID=1496;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=VPI 10463;
 RX MEDLINE=90326540; PubMed=2374729;
 RA Barrero L.A., Wang S.Z., Phelps C.J., Johnson U.J., Wilkins T.D.,
 RT "Nucleotide sequence of Clostridium difficile toxin B gene."
 RL Nucleic Acids Res. 18:4004-4004(1990).

RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=VPI 10463;
 RA von Bichel-Streiber C.;
 RN Submitted (JAN-1997) to the EMBL/Genbank/DBJ databases.
 RN [3]
 RP SEQUENCE OF 1271-2366 FROM N.A.
 RC STRAIN=VPI 10463;
 RX MEDLINE=92293124; PubMed=1603066;
 RA Bichel-Streiber C., Laufenberg-Feldmann R., Sarlingen S., Schlze J.,
 RA Sauerborn M.,
 RT "Comparative sequence analysis of the Clostridium difficile toxins A
 RT and B."
 RL Mol. Gen. Genet. 233:260-268(1992).
 CC -1- DISEASE: CLOSTRIDIUM DIFFICILE PRODUCES TWO TOXINS, AN ENTEROTOXIN
 CC CALLED A AND CYTOTOXIN B.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; X53138; CAA37298.1; -
 DR EMBL; X92982; CAA63562.1; -
 DR PIR; X60984; CAA43299.1; -
 DR PIR; A27636; A27636.
 DR PIR; S10317; S10317.
 DR InterPro; IPR002479; CW binding.
 DR Pfam; PF01473; CW binding 1; 17.
 DR Pfam; PF04488; Gly_transf_sug; 1.
 KW Toxin.
 SQ SEQUENCE 2366 AA; 269709 MW; E1024BD8B8A56ADF CRC64;
 QY
 Query Match 39.8%; Score 45; DB 1; Length 2366;
 Best Local Similarity 44.4%; Pred. No. 1.1e+02;
 Matches 8; Conservative 4; Mismatches 6; Indels 0; Gaps 0;
 Db 1733 INDLIRYWSNDGNDPI 1750
 2 IDHLSILEAWSGNDNDYV 19
 RESULT 12
 ACOB_BACSU
 ID ACOB_BACSU STANDARD; PRT; 341 AA.
 AC O34591;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Acetoin:2,6-dichlorophenolindophenol oxidoreductase beta subunit
 DE (EC 1.1.1.-) (Acetoin:DCPIP oxidoreductase-beta) (AO:DCPIP OR)
 DE (TPP-dependent acetoin dehydrogenase E1 beta-subunit).
 GN ACOB.
 OS Bacillus subtilis.
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 OX NCBI_TaxID=1423;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=168;
 RX MEDLINE=99296597; PubMed=10368162;
 RA Huang M., Oppermann-Sanio F.B., Steibuchel A.,
 RT "Biochemical and molecular characterization of the Bacillus subtilis
 RT acetoin catabolic pathway."
 RL J. Bacteriol. 181:3837-3841(1999).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=AC327;
 RX MEDLINE=97124190; PubMed=8969503;
 RA Yamamoto H., Uchiyama S., Sekiguchi J.,
 RT "Cloning and sequencing of a 40.6 kb segment in the 73 degrees-76

RT degrees region of the Bacillus subtilis chromosome containing genes
 RT for trehalose metabolism and acetoin utilization.";
 RT Microbiology 142:3057-3065(1996).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=168;
 RX MEDLINE=98044033; PubMed=9384377;
 RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
 RA Azevedo V., Bertero M.G., Besieres P., Bolotin A., Borchert S.,
 RA Bortis R., Boursier L., Brans A., Braun M., Brignelli S.C., Bron S.,
 RA Brouillet S., Brusch C.V., Caldwell B., Capuano V., Carter N.M.,
 RA Choi S.K., Codani J.J., Conterton I.F., Cummings N.J., Daniel R.A.,
 RA Denzot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emerson P.T.,
 RA Entian K.D., Erington J., Fabret C., Ferrati E., Fougere D.,
 RA Frits C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,
 RA Ghim S.Y., Glaeser P., Goffeau A., Golightly E.J., Grandi G.,
 RA Giuseppe G., Guy B.J., Haga K., Harech J., Harwood C.R., Henaut A.,
 RA Hilbert H., Holstappel S., Hosono S., Hullo M.F., Ilaya M., Jones L.,
 RA Joris B., Karamata D., Kasahara Y., Klaerr-Bianchand M., Klein C.,
 RA Kobayashi Y., Koeltter P., Koningsstein G., Krogh S., Kumano M.,
 RA Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,
 RA Lee S.M., Levine A., Liu H., Masuda S., Manuel C., Medigue C.,
 RA Medina N., Mellado R.P., Mizuno M., Moesti D., Nakai S., Noback M.,
 RA Noone D., O'Reilly M., Ogawa K., Ogawa K., Oudega B., Park S.H.,
 RA Parro V., Pohl T.M., Portetelle D., Rapoport G., Prescott A.M.,
 RA Pirescan E., Puig C., Purnelle B., Rapoport G., Rey M., Reynolds S.,
 RA Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadale Y.,
 RA Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,
 RA Sekiguchi J., Sekowska A., Seror S.J., Seror P., Shin B.S., Soldo B.,
 RA Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,
 RA Takeuchi M., Yamakoshi A., Tanaka T., Terpstra P., Tognoni A.,
 RA Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassartoli A.,
 RA Viari A., Wamput R., Wedler H., Wedler H., Wellenreger T.,
 RA Winers P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
 RA Yoshida K., Yoshikawa H.F., Zumschein E., Yoshikawa H., Danchin A.,
 RT "The complete genome sequence of the Gram-positive bacterium Bacillus
 RT subtilis".
 RT Nature 390:249-256(1997).
 CC -1- FUNCTION: CATALYZES THE 2,6-DICHLOROPHENOLINDOPHENOL-DEPENDENT
 CC CLEAVAGE OF ACETONIN INTO ACETATE AND ACETALDEHYDE (BY SIMILARITY).
 CC -1- SUBSTRATE: TETRAMER OF TWO ALPHA AND TWO BETA SUBUNITS (BY
 CC SIMILARITY).
 CC -1- SIMILARITY: TO THE BETA SUBUNITS OF 2-OXO-ACID DEHYDROGENASE
 CC COMPONENTS OF VARIOUS MULTISUBUNIT COMPLEXES.
 CC CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC DR EMBL; AF006075; AAC05583.1; -
 CC DR EMBL; D78509; BAA24295.1; -
 CC DR EMBL; Z99108; CAB12636.1; -
 CC DR PIR; B69581; B69581.
 CC DR HSSP; P09061; 1080.
 CC DR Subtilist; BG12559; accB.
 CC DR InterPro; IPR005476; Transketolase C.
 CC DR InterPro; IPR005475; Transketolase CR.
 CC DR Pfam; PF02779; transket pyr; 1.
 CC DR Pfam; PF02780; transketolase_C; 1.
 CC DR Acetoin catabolism; Oxidoreductase; Complete proteome.
 CC FT INIT MET 0
 CC BY SIMILARITY.
 CC SQ SEQUENCE 341 AA; 36713 MM; DCSB3D549C89EFO CRC64;
 CC
 CC Query Match 38.9%; Score 44; DB 1; Length 341;
 CC Best Local Similarity 61.5%; Pred. No. 17;
 CC Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Db 36 AYDHLQDDEAWG 48
 ||| |||
 RESULT 13
 ID STK4_HUMAN STANDARD; PRT; 487 AA.
 AC Q13043; Q15802; Q9NT24;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Serine/threonine protein kinase 4 (EC 2.7.1.37) (STE20-like kinase
 DE MST1) (MST-1) (Mammalian STE20-like protein kinase 1)
 DE (Serine/threonine protein kinase Krs-2).
 OS STK4 OR MST1.
 GN Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=95394929; PubMed=7665586;
 RA Creasy C.L., Chernoff J.;
 RT "Cloning and characterization of a human protein kinase with homology
 RT to Ste20".
 RT J. Biol. Chem. 270:21695-21700(1995).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96413604; PubMed=8816758;
 RA Taylor L.K., Wang H.C., Erikson R.L.;
 RT "Newly identified stress-responsive protein kinases, Krs-1 and Krs-
 RT 2".
 RT Proc. Natl. Acad. Sci. U.S.A. 93:10099-10104(1996).
 RN [3]
 RP SEQUENCE OF 1-435 FROM N.A.
 RL Laird G.;
 CC Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 CC -1- CATALYTIC ACTIVITY: ADP + a protein = ADP + a phosphorylated protein.
 CC -1- ENZYME REGULATION: THE C-TERMINAL NON-CATALYTIC REGION INHIBITS
 CC THE KINASE ACTIVITY.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 CC -1- TISSUE SPECIFICITY: Ubiquitously expressed.
 CC -1- PM: AUTOPHOSPHORYLATED ON SERINE AND THREONINE RESIDUES.
 CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 CC STE20 SUBFAMILY.
 CC CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC DR EMBL; U18297; AAA83254.1; -
 CC DR EMBL; U60207; AAB17262.1; -
 CC DR EMBL; AL109839; CAB89421.1; -
 CC DR HSSP; P24941; IHCL.
 CC DR Genew; HGNC:11408; STK4.
 CC DR MIM; 604965; -
 CC DR GO; GO:0004674; F:protein serine/threonine kinase activity; TAS.
 CC DR GO; GO:0006468; P:protein amino acid phosphorylation; TAS.
 CC DR GO; GO:0007165; P:signal transduction; TAS.
 CC DR InterPro; IPR000719; Prot_kinase.
 CC DR InterPro; IPR002290; Ser_thr_kinase.
 CC DR InterPro; IPR001245; Tyr_kinase.
 CC DR Pfam; PF00069; pkinase; 1.
 CC DR PRINTS; PR00109; TYRKINASE.
 CC DR PRODOM; PD000001; Prot_kinase; 1.
 CC DR SMART; SM00220; S_TKc; 1.
 CC DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
 CC DR PROSITE; PS00108; PROTEIN_KINASE_ST; FALSE_NEG.
 CC DR PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.

Transferase; Serine/threonine-protein kinase; ATP-binding.
 KM DOMAIN 30 281 PROTEIN KINASE.
 FT NP BIND 36 44 ATP (BY SIMILARITY).
 FT BINDING 59 59 ATP (BY SIMILARITY).
 FT ACT_SITE 149 149 BY SIMILARITY.
 FT DOMAIN 373 378 POLY-GLU.
 FT CONFLICT 222 222 P -> R (IN REF. 1).
 FT CONFLICT 312 312 V -> M (IN REF. 1).
 SO SEQUENCE 487 AA; 55630 MW; 150758B5C5F77D5C CRC64;

Query Match 38.9%; Score 44; DB 1; Length 487;
 Best Local Similarity 61.5%; Pred. No. 26;
 Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 9 EAMSGNDNDYKQ 21
 Db 248 ELWSDNFTDFVKQ 260

RESULT 14
 ID PGMT_DICDI STANDARD; PRT; 572 AA.
 AC Q23919;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Phosphoglucosyltransferase (EC 5.4.2.2) (Glucose phosphotransferase) (PGM).
 GN PGM.
 OS Dictyostelium discoideum (Slime mold).
 OC Eukaryota; Mycetozoa; Dictyostelidae; Dictyostelium.
 OX NCBI_TaxID=44689;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=AX4;
 RA Loomis W.F.;
 RL Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: THIS ENZYME PARTICIPATES IN BOTH THE BREAKDOWN AND
 CC SYNTHESIS OF GLUCOSE (BY SIMILARITY).
 CC -1- CATALYTIC ACTIVITY: Alpha-D-glucose 1-phosphate = alpha-D-glucose
 CC + 6-phosphate.
 CC -1- COFACTOR: REQUIRES MAGNESIUM FOR ACTIVITY (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE PHOSPHOHEXOSE MUTASES FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC
 CC EMBL; U61984; AAB03667.1; -.
 DR HSSP; P00949; 3PMG.
 DR DictyDb; DD05092; PGM.
 DR InterPro; IPR005841; PG/PMM mutase.
 DR InterPro; IPR005844; PG_PMM_ABAT.
 DR InterPro; IPR005845; PG_PMM_ABAT.
 DR InterPro; IPR005846; PG_PMM_ABAT.
 DR InterPro; IPR005843; PG_PMM_C.
 DR Pfam; PF00408; PGM_PMM_1.
 DR Pfam; PF02878; PGM_PMM_1.
 DR Pfam; PF02879; PGM_PMM_1.
 DR Pfam; PF02880; PGM_PMM_1.
 DR PRINTS; PR00509; PGM_PMM.
 DR PROSITE; PS00710; PGM_PMM_1.
 DR Isoemase; Phosphorylation; Magnesium.
 KW ACT_SITE 120 120 PHOSPHOSERINE INTERMEDIATE
 FT ACT_SITE 120 120 (BY SIMILARITY).
 SO SEQUENCE 572 AA; 63229 MW; ED7E8F58F16F563A CRC64;

Query Match 38.1%; Score 43; DB 1; Length 572;
 Best Local Similarity 31.6%; Pred. No. 44;

Matches 6; Conservative 8; Mismatches 5; Indels 0; Gaps 0;

Qy 1 AIDHSLILAMSGNDNDYV 19
 Db 167 SYDNLGLTKTYEWNDDGEFV 185

RESULT 15
 ID COA_BACTU STANDARD; PRT; 674 AA.
 AC O87905;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE pesticidal crystal protein cry24Aa (insecticidal delta-endotoxin
 DE CryxxIva(a)) (crystalline entomocidal protoxin) (Crystal protein)
 DE (insecticidal protein Jegg72) (Fragment).
 GN CRY24AA OR CRYXXIVA(A).
 OS Bacillus thuringiensis (subsp. jegashtesan).
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 OX NCBI_TaxID=56955;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Kawalek M.D., Gill S.S.;
 RT "Isolation and characterization of insecticidal genes from Bacillus
 RT thuringiensis subsp. jegashtesan."
 RL Submitted (JAN-1997) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: PROMOTES COLLOIDSMOTIC LYSIS BY BINDING TO THE MIDGUT
 CC EPITHELIAL CELLS OF INSECTS.
 CC -1- DEVELOPMENTAL STAGE: THE CRYSTAL PROTEIN IS PRODUCED DURING
 CC SPOULATION AND IS ACCUMULATED BOTH AS AN INCLUSION AND AS PART
 CC OF THE SPORE COAT.
 CC -1- MISCELLANEOUS: TOXIC SEGMENT OF THE PROTEIN IS LOCATED IN THE
 CC N-TERMINUS.
 CC -1- SIMILARITY: BELONGS TO THE DELTA ENDOTOXIN FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC
 CC EMBL; U86188; AAC61891.1; -.
 DR InterPro; IPR001178; Endotoxin.
 DR InterPro; IPR005638; endotoxin_C.
 DR InterPro; IPR005639; endotoxin_N.
 DR Pfam; PF00555; endotoxin_1.
 DR Pfam; PF03944; endotoxin_C; 1.
 DR Pfam; PF03945; endotoxin_N; 1.
 DR Toxin; Sporulation.
 KW NON_TER 674 674
 FT NON_TER 674 674
 SO SEQUENCE 674 AA; 75959 MW; DA3904DAB891C978 CRC64;

Query Match 38.1%; Score 43; DB 1; Length 674;
 Best Local Similarity 46.7%; Pred. No. 53;
 Matches 7; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

Qy 3 DRLSTLEAMSGNDND 17
 Db 138 DYLGALEAMNNKSN 152

Search completed: November 13, 2003, 09:45:32
 Job time: 8.56398 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 13, 2003, 09:31:40 ; Search time 31.6493 Seconds
(without alignments)
171.224 Million cell updates/sec

Title: US-09-290-049A-13
Perfect score: 113
Sequence: 1 AIDHLSILEAMSGNDNDYVKQ 21

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 830525 seqs, 258052604 residues
Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

SPTREMBL_23:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phage:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_virus:*
- 16: sp_bacteriap:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	108	95.6	1338	2 Q9WJ4	Q9WJ4 streptococc
2	79	69.9	1575	2 Q9LCH3	Q9LCH3 streptococc
3	77	68.1	1577	2 Q54178	Q54178 streptococc
4	73	64.6	1572	2 Q55265	Q55265 streptococc
5	71	62.8	1554	2 Q8KZL5	Q8KZL5 streptococc
6	69	61.1	1477	2 Q9L466	Q9L466 streptococc
7	69	61.1	1508	2 Q9E2H5	Q9E2H5 leuconostoc
8	69	61.1	1508	2 Q52224	Q52224 leuconostoc
9	69	61.1	1512	2 Q9WJX5	Q9WJX5 streptococc
10	69	61.1	1518	2 Q00600	Q00600 streptococc
11	67	59.3	1449	2 Q06842	Q06842 streptococc
12	67	59.3	1449	2 Q55264	Q55264 streptococc
13	66	58.4	1016	2 Q9LCT7	Q9LCT7 leuconostoc
14	65	57.5	2057	2 Q9RB05	Q9RB05 leuconostoc
15	64	56.6	2835	2 Q8G9Q2	Q8G9Q2 leuconostoc
16	62	54.9	1590	2 Q59983	Q59983 streptococc

17	62	54.9	1590	2 Q55263	Q55263 streptococc
18	61	54.0	1599	2 Q00599	Q00599 streptococc
19	58	51.3	1527	2 Q9ZAR4	Q9ZAR4 leuconostoc
20	58	51.3	1527	2 Q8KRE1	Q8KRE1 leuconostoc
21	49	43.4	1428	17 Q9HSM9	Q9HSM9 halobacteri
22	49	43.4	1428	3 Q8TG58	Q8TG58 saccharomyc
23	49	43.4	1428	3 Q8TG46	Q8TG46 saccharomyc
24	49	43.4	1428	3 Q8TG44	Q8TG44 saccharomyc
25	49	43.4	1428	3 Q8TG53	Q8TG53 saccharomyc
26	49	43.4	1428	3 Q8TG56	Q8TG56 saccharomyc
27	49	43.4	1428	3 Q8TR86	Q8TR86 saccharomyc
28	49	43.4	1428	3 Q8TG47	Q8TG47 saccharomyc
29	49	43.4	1428	3 Q8TG43	Q8TG43 saccharomyc
30	48.5	42.9	576	4 Q96DX4	Q96DX4 homo sapien
31	48.5	42.9	576	6 Q95LP3	Q95LP3 macaca fasc
32	48	42.5	378	10 Q9SZH3	Q9SZH3 arabidopsi
33	48	42.5	382	16 Q8UJ06	Q8UJ06 agrobacteri
34	48	42.5	394	10 Q9AKJ2	Q9AKJ2 arabidopsi
35	48	42.5	695	16 Q9Z9G1	Q9Z9G1 chlamydia p
36	48	42.5	704	16 Q9KIZ6	Q9KIZ6 chlamydia p
37	48	42.5	956	4 Q8NIS3	Q8NIS3 homo sapien
38	48	42.5	3198	12 Q9IW34	Q9IW34 pea seed-bo
39	48	42.5	6885	4 Q8WXH0	Q8WXH0 homo sapien
40	47	41.6	298	5 Q19058	Q19058 caenorhabdi
41	47	41.6	769	10 Q8LAE1	Q8LAE1 arabidopsi
42	47	41.6	769	10 Q9LNU1	Q9LNU1 arabidopsi
43	47	41.6	3199	12 Q85074	Q85074 pea seed-bo
44	46	40.7	255	11 Q9CXL4	Q9CXL4 mus muscula
45	46	40.7	287	16 Q9RK11	Q9RK11 streptomyce

ALIGNMENTS

RESULT 1

Q9WJ4 PRELIMINARY; PRT; 1338 AA.

AC Q9WJ4;
DT 01-NOV-1999 (TREMBlrel. 12, Created)
DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
DE GTF-S.
GN GTF-S.
OS Streptococcus criceti.
OC Plasmid PAM.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1333;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HS-6.
RA Inoue M., Fukui K., Miyagi A.;
RT "S.cricetus glucosyltransferase(gtfS and gtfT) genes."
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AB026123; BAA77236.1; -
DR InterPro: IPR002479; CW_binding.
DR InterPro: IPR003318; Glyco_hydro_70.
DR Pfam: PF01473; CW_binding_1; 10.
DR Pfam: PF02324; Glyco_hydro_70; 1.
KW Plasmid.
SQ
SEQUENCE 1338 AA; 148558 MW; 0A90C8E10E15D99B CRC64;

Query Match 95.6%; Score 108; DB 2; Length 1338;
Best Local Similarity 100.0%; Pred. No. 46-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AIDHLSILEAMSGNDNDYVK 20
Db 437 AIDHLSILEAMSGNDNDYVK 456

RESULT 2
Q9LCH3

```
ID 091CH3 PRELIMINARY; PRT; 1575 AA.
AC Q9LCH3;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Glucosyltransferase.
GN GTFP.
OS Streptococcus oralis.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1303;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC10557;
RX MEDLINE=20231779; PubMed=10768934;
RA Fujiwara T., Hoshino T., Ooshima T., Sobue S., Hamada S.;
RT "Purification, characterization, and molecular analysis of the gene
  encoding glucosyltransferase from Streptococcus oralis.";
RL Infect. Immun. 68:2475-2483(2000).
DR EMBL: AB025228; BA95201.1; -;
DR InterPro: IPR002479; CW_binding.
DR InterPro: IPR003318; Glyco_hydro_70.
DR Pfam: PF01473; CW_binding.1; 17.
DR Pfam: PF02324; Glyco_hydro_70; 1.
KW Transferase.
SQ SEQUENCE 1575 AA; 176792 MM; 772A26E4D7C2E543 CRC64;

Query Match 69.9%; Score 79; DB 2; Length 1575;
Best Local Similarity 80.0%; Pred. No. 0.0019;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AIDHLSILEAWSGNDNDYK 20
DB 546 AIKHLSTLEAWSDNDPDYK 565

RESULT 3
ID 054178 PRELIMINARY; PRT; 1577 AA.
AC Q54178; Q54247;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Glucosyltransferase.
GN GTFG.
OS Streptococcus gordonii Challis.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=29390;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CHALLIS;
RX MEDLINE=96157084; PubMed=8586195;
RA Vickerman M.M., Sulavik M.C., Clewell D.B.;
RT "Molecular analysis of Streptococcus gordonii glucosyltransferase
  phase variants.";
RL Dev. Biol. Stand. 85:309-314(1995).
RN [2]
RP SEQUENCE OF 1-96 FROM N.A.
RC STRAIN=CHALLIS;
RX MEDLINE=92276337; PubMed=1534326;
RA "Identification of a gene, tsg, which regulates expression of
  glucosyltransferase and influences the Spp phenotype of Streptococcus
  gordonii Challis.";
RL J. Bacteriol. 174:3577-3586(1992).
DR EMBL: U12643; AAC43483.1; -;
DR EMBL: M89776; AAA26969.1; -;
DR InterPro: IPR002479; CW_binding.
DR InterPro: IPR003318; Glyco_hydro_70.
DR Pfam: PF01473; CW_binding.1; 18.
DR Pfam: PF02324; Glyco_hydro_70; 1.
KW Transferase.
RT
```

```
SQ SEQUENCE 1577 AA; 177805 MM; 5AE0328DC5E08D18 CRC64;

Query Match 68.1%; Score 77; DB 2; Length 1577;
Best Local Similarity 75.0%; Pred. No. 0.004;
Matches 15; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 AIDHLSILEAWSGNDNDYK 20
DB 548 AIKHLSTLEAWSDNDPDYK 567

RESULT 4
ID 055265 PRELIMINARY; PRT; 1577 AA.
AC Q55265;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Glucosyltransferase precursor.
GN GTFM.
OS Streptococcus salivarius.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1304;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95122197; PubMed=7822030;
RA Simpson C.L., Giffard P.M., Jacques N.A.;
RT "Streptococcus salivarius ATCC 25975 possesses at least two genes
  coding for primer-independent glucosyltransferases.";
RL Infect. Immun. 63:609-621(1995).
DR EMBL: J35928; AAC41413.1; -;
DR InterPro: IPR004829; Ceuface_antigen.
DR InterPro: IPR002479; CW_binding.
DR InterPro: IPR003318; Glyco_hydro_70.
DR Pfam: PF01473; CW_binding.1; 11.
DR Pfam: PF02324; Glyco_hydro_70; 1.
DR Pfam: PD153432; Ceuface_antigen; 1.
KW Signal; Transferase.
FT SIGNAL 1 38 POTENTIAL.
FT CHAIN 39 1577 GLUCOSYLTRANSFERASE.
SQ SEQUENCE 1577 AA; 175290 MM; 3EFB898A7D3A7BE3 CRC64;

Query Match 64.6%; Score 73; DB 2; Length 1577;
Best Local Similarity 75.0%; Pred. No. 0.017;
Matches 15; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 AIDHLSILEAWSGNDNDYK 20
DB 588 AIKHLSTLEAWSYNDHQYK 607

RESULT 5
ID 08KZL5 PRELIMINARY; PRT; 1554 AA.
AC 08KZL5;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Glucosyltransferase.
GN GTFU.
OS Streptococcus sobrinus.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1310;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21958684; PubMed=11960691;
RA Hanada N., Fukushima K., Nomura Y., Senguku H., Hayakawa M.,
  Mukasa H., Shiroza T., Abiko Y.;
RT "Cloning and nucleotide sequence analysis of the Streptococcus
  sobrinus gltf gene that produces a highly branched water-soluble
  glucan.";
RT
```

RL Bioclim. Biophys. Acta 1570:75-79 (2002).
DR EMBL; AB089438; BAC07265.1; -.
DR InterPro; IPR002479; CW_binding.
DR InterPro; IPR003318; Glyco_hydro_70.
DR Pfam; PF01473; CW_binding_1; 6.
DR Pfam; PF02324; Glyco_hydro_70; 1.
KM Transferase.
SQ SEQUENCE 1554 AA; 171676 MW; 6981BCC1DAE24A73 CRC64;
Query Match 62.8%; Score 71; DB 2; Length 1554;
Best Local Similarity 72.2%; Pred. No. 0.035;
Matches 13; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
QY 1 AIDHLSILEAMSGNDNDYK 18
Db 485 ALAHISILEAMSLDNDQY 502
RESULT 6
ID 091466 PRELIMINARY; PRT; 1477 AA.
AC 091466;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
DE Dextranucrase (EC 2.4.1.5).
GN DSRG.
OS Leuconostoc mesenteroides.
OC Bacteria; Firmicutes; Lactobacillales; Leuconostoc.
OX NCBI_TaxID=1245;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NRRL B-1355;
RA Arguello-Morales M.A., Remaud-Simeon M., Pizzut S., Sarcabal P., Willemot R.M., Monsan P.;
RT "Sequence analysis of the gene encoding alternanucrase, a sucrose glucosyltransferase from Leuconostoc mesenteroides NRRL B-1355.";
RT Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
RL EMBL; AJ250172; CB876565.1; -.
DR InterPro; IPR002479; CW_binding.
DR InterPro; IPR003318; Glyco_hydro_70.
DR Pfam; PF01473; CW_binding_1; 14.
DR Pfam; PF02324; Glyco_hydro_70; 1.
KM Glycosyltransferase; Transferase.
SQ SEQUENCE 1477 AA; 164887 MW; B6F5710DEDFC831 CRC64;
Query Match 61.1%; Score 69; DB 2; Length 1477;
Best Local Similarity 70.0%; Pred. No. 0.069;
Matches 14; Conservative 1; Mismatches 5; Indels 0; Gaps 0;
QY 1 AIDHLSILEAMSGNDNDYK 20
Db 532 ANQHLSTLEWSHNDPEYK 551
RESULT 7
ID 09EZHS PRELIMINARY; PRT; 1508 AA.
AC 09EZHS;
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
DE Dextranucrase Derb742.
GN DSRB742.
OS Leuconostoc mesenteroides.
OC Bacteria; Firmicutes; Lactobacillales; Leuconostoc.
OX NCBI_TaxID=1245;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=B-742CB;
RA Kim H.-S., Kim D., Ryu H.-J., Robyt J.F.;
RT "Leuconostoc mesenteroides B-742CB, a dextranucrase gene,"
Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.

DR EMBL; AF294469; AAG38021.1; -.
DR InterPro; IPR002479; CW_binding.
DR InterPro; IPR003318; Glyco_hydro_70.
DR Pfam; PF01473; CW_binding_1; 14.
DR Pfam; PF02324; Glyco_hydro_70; 1.
SQ SEQUENCE 1508 AA; 168542 MW; E2FCFA087A8AFA2A CRC64;
Query Match 61.1%; Score 69; DB 2; Length 1508;
Best Local Similarity 70.0%; Pred. No. 0.07;
Matches 14; Conservative 1; Mismatches 5; Indels 0; Gaps 0;
QY 1 AIDHLSILEAMSGNDNDYK 20
Db 563 ANQHLSTLEWSHNDPEYK 582
RESULT 8
ID 052224 PRELIMINARY; PRT; 1508 AA.
AC 052224;
DT 01-JUN-1998 (TREMBlrel. 06, Created)
DT 01-JUN-1998 (TREMBlrel. 06, Last sequence update)
DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
DE Glucosyltransferase (EC 2.4.1.5).
GN DSRB.
OS Leuconostoc mesenteroides.
OC Bacteria; Firmicutes; Lactobacillales; Leuconostoc.
OX NCBI_TaxID=1245;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NRRL B-1299;
RA Monchois V., Remaud-Simeon M., Monsan P., Willemot R.M.;
RT "Cloning and sequencing of a gene coding for an extracellular dextranucrase (DSRB) from Leuconostoc mesenteroides NRRL B-1299
RT synthesizing only a (1-6) glucan.";
RT FEMS Microbiol. Lett. 0:0-0(1998).
RL EMBL; AF030129; AAB95453.1; -.
DR InterPro; IPR002479; CW_binding.
DR InterPro; IPR003318; Glyco_hydro_70.
DR Pfam; PF01473; CW_binding_1; 14.
DR Pfam; PF02324; Glyco_hydro_70; 1.
KM Glycosyltransferase; Transferase.
SQ SEQUENCE 1508 AA; 168511 MW; E70CCE857A70D1F0 CRC64;
Query Match 61.1%; Score 69; DB 2; Length 1508;
Best Local Similarity 70.0%; Pred. No. 0.07;
Matches 14; Conservative 1; Mismatches 5; Indels 0; Gaps 0;
QY 1 AIDHLSILEAMSGNDNDYK 20
Db 563 ANQHLSTLEWSHNDPEYK 582
RESULT 9
ID 09WKJ5 PRELIMINARY; PRT; 1512 AA.
AC 09WKJ5;
DT 01-NOV-1999 (TREMBlrel. 12, Created)
DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
DE GTF-S.
GN GTF.
OS Streptococcus criceti.
OC Plasmid pAM1.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1333;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HS-6;
RA Inoue M., Fukui K., Miyagi A.;
RT "S. cricetus glucosyltransferase(gtfS and gtfT) genes,"
Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.

DR EMBL; AB026123; BAA77237.1; -
DR InterPro; IPR002479; CM_binding.
DR InterPro; IPR00318; Glyco_hydro_70.
DR Pfam; PF01473; CM_binding_1; 14.
DR Pfam; PF02324; Glyco_hydro_70; 1.
KW Pfamid.
SQ SEQUENCE 1512 AA; 167145 MW; 4C03D9C6C01FC14 CRC64;

Query Match 61.1%; Score 69; DB 2; Length 1512;
Best Local Similarity 72.2%; Pred. No. 0.071;
Matches 13; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 AIDHLSILEAMSGNDNDY 18
|:|||||
Db 485 ALAHISILEAMSGNDNDY 502

RESULT 10
ID 000600 PRELIMINARY; PRT; 1518 AA.
AC 000600;
DT 01-NOV-1996 (TRENBLREL. 01, Created)
DT 01-NOV-1996 (TRENBLREL. 01, Last sequence update)
DT 01-OCT-2002 (TRENBLREL. 22, Last annotation update)
DE Glucosyltransferase I (EC 2.4.1.5) (GTF) (Dextranucrase) (Sucrose 6-
DE Glucosyltransferase).
GN GTF.
OS Streptococcus salivarius.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1304;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 25975;
RX MEDLINE=92148377; PubMed=1838391;
RA Giffard P.M., Simpson C.L., Milward C.P., Jacques N.A.;
RT "Molecular characterization of a cluster of at least two
RT glucosyltransferase genes in Streptococcus salivarius ATCC 25975.";
RL J. Gen. Microbiol. 137:2577-2593 (1991).
CC -1- FUNCTION: PRODUCTION OF EXTRACELLULAR GLUCANS, THAT ARE THOUGHT TO
CC PLAY A KEY ROLE IN THE DEVELOPMENT OF THE DENTAL PLAQUE BECAUSE OF
CC THEIR ABILITY TO ADHERE TO SMOOTH SURFACES AND MEDIATE THE
CC AGGREGATION OF BACTERIAL CELLS AND FOOD DEBRIS.
CC -1- CATALYTIC ACTIVITY: SUCROSE + ((1,6)-ALPHA-D-GLUCOSYL) (N) = D-
CC FRUCTOSE + ((1,6)-ALPHA-D-GLUCOSYL) (N+1).
CC -1- SUBCELLULAR LOCATION: EXTRACELLULAR.
CC -1- DISEASE: DENTAL CARIES.
DR EMBL; Z11873; CAAT7900.1; -
DR EMBL; M61111; AA26896.1; -
DR InterPro; IPR002479; CM_binding.
DR InterPro; IPR00318; Glyco_hydro_70.
DR Pfam; PF01473; CM_binding_1; 13.
DR Pfam; PF02324; Glyco_hydro_70; 1.
KW Transferrase, Glucosyltransferase; Repeat; Dental caries.
FT DOMAIN 1307 1482 6 DIRECT REPEATS.
FT REPEAT 1307 1338 REPEAT 1.
FT REPEAT 1338 1352 REPEAT 2.
FT REPEAT 1372 1403 REPEAT 3.
FT REPEAT 1404 1417 REPEAT 4.
FT REPEAT 1437 1468 REPEAT 5.
FT REPEAT 1469 1482 REPEAT 6.
SQ SEQUENCE 1518 AA; 167730 MW; DAA41F717098B59A CRC64;

Query Match 61.1%; Score 69; DB 2; Length 1518;
Best Local Similarity 66.7%; Pred. No. 0.071;
Matches 12; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 AIDHLSILEAMSGNDNDY 18
|:|||||
Db 501 ALAHISILEAMSGNDNDY 518

RESULT 11

068542
ID 068542 PRELIMINARY; PRT; 1449 AA.
AC 068542;
DT 01-AUG-1998 (TRENBLREL. 07, Created)
DT 01-AUG-1998 (TRENBLREL. 07, Last sequence update)
DT 01-OCT-2002 (TRENBLREL. 22, Last annotation update)
DE Glucosyltransferase N (Fragment).
GN GTFN.
OS Streptococcus salivarius.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1304;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=V1477;
RA Jaffe R.I.;
RT "Streptococcus salivarius V1477 gtfN".
RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF049609; AAC05156.1; -
DR InterPro; IPR002479; CM_binding.
DR InterPro; IPR00318; Glyco_hydro_70.
DR Pfam; PF01473; CM_binding_1; 8.
DR Pfam; PF02324; Glyco_hydro_70; 1.
KW Transferrase.
FT NON-TER
SQ SEQUENCE 1449 AA; 159895 MW; 0700F6D748471BFB CRC64;

Query Match 59.3%; Score 67; DB 2; Length 1449;
Best Local Similarity 77.8%; Pred. No. 0.14;
Matches 14; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AIDHLSILEAMSGNDNDY 18
|:|||||
Db 536 AIKHLISILEAMSGNDNDY 553

RESULT 12
ID 055264 PRELIMINARY; PRT; 1449 AA.
AC 055264;
DT 01-NOV-1996 (TRENBLREL. 01, Created)
DT 01-NOV-1996 (TRENBLREL. 01, Last sequence update)
DT 01-OCT-2002 (TRENBLREL. 22, Last annotation update)
DE Glucosyltransferase precursor.
GN GTF.
OS Streptococcus salivarius.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1304;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95122197; PubMed=7822030;
RA Simpson C.L., Giffard P.M., Jacques N.A.;
RT "Streptococcus salivarius ATCC 25975 possesses at least two genes
RT coding for primer-independent glucosyltransferases.";
RL Infect. Immun. 63:609-621 (1995).
DR EMBL; J35495; AAC41412.1; -
DR InterPro; IPR002479; CM_binding.
DR InterPro; IPR00318; Glyco_hydro_70.
DR Pfam; PF01473; CM_binding_1; 8.
DR Pfam; PF02324; Glyco_hydro_70; 1.
KW Signal; Transferrase.
FT SIGNAL 1 35
FT CHAIN 36 1449 POTENTIAL.
SQ SEQUENCE 1449 AA; 159984 MW; DD62F07306E86A46 CRC64;

Query Match 59.3%; Score 67; DB 2; Length 1449;
Best Local Similarity 77.8%; Pred. No. 0.14;
Matches 14; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AIDHLSILEAMSGNDNDY 18
|:|||||
Db 536 AIKHLISILEAMSGNDNDY 553

RESULT 13
09LCJ7
ID 09LCJ7 PRELIMINARY; PRT; 1016 AA.
AC 09LCJ7;
DT 01-OCT-2000 (TRENBLREL. 15, Created)
DT 01-OCT-2000 (TRENBLREL. 15, Last sequence update)
DT 01-OCT-2002 (TRENBLREL. 22, Last annotation update)
DE Dextranucrase.
GN DSRT.
OS Leuconostoc mesenteroides.
OC Bacteria; Firmicutes; Lactobacillales; Leuconostoc.
OX NCBI_TaxID=1245;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NRRL B-512F;
RX MEDLINE=20169623; PubMed=10705445;
RA Funane K., Mizuno K., Takahara H., Kobayashi M.;
RT "Gene encoding a dextranucrase-like protein in Leuconostoc
RL mesenteroides NRRL B-512F.", 64:29-38(2000).
DR BMBJ; Biotechnol. Biochem. 64:29-38(2000).
DR HSSP; P06278; IVOS.
DR InterPro; IPR003318; Glyco_hydro.70.
DR Pfam; PF02324; Glyco_hydro.70; 1_
SQ SEQUENCE 1016 AA; 110344 MW; 8896EFDE13CCB47 CRC64;

Query Match 58.4%; Score 66; DB 2; Length 1016;
Best Local Similarity 70.6%; Pred. No. 0.13;
Matches 12; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 4 HLSTLEAWSGNDNDYVK 20
Db 551 HVSILEDWSDNDADRYK 567

RESULT 14
09RE05
ID 09RE05 PRELIMINARY; PRT; 2057 AA.
AC 09RE05;
DT 01-MAY-2000 (TRENBLREL. 13, Created)
DT 01-MAR-2001 (TRENBLREL. 16, Last sequence update)
DT 01-OCT-2002 (TRENBLREL. 22, Last annotation update)
DE Alernansucrase (EC 2.4.1.140).
GN ASR.
OS Leuconostoc mesenteroides.
OC Bacteria; Firmicutes; Lactobacillales; Leuconostoc.
OX NCBI_TaxID=1245;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NRRL B-1355;
RX MEDLINE=20080809; PubMed=10612736;
RA Arzuellio-Morales M.A., Remaud-Simeon M., Pizzut S., Sarcabal P.,
RA Willemot R.M., Monsan P.;
RT "Sequence analysis of the gene encoding alernansucrase, a sucrose
RT glucosyltransferase from Leuconostoc mesenteroides NRRL B-1355.",
RL FEMS Microbiol. Lett. 182:81-85(2000).
DR EMBL; AJ250173; CAB55910.2;
DR InterPro; IPR002479; CM binding.
DR InterPro; IPR003318; Glyco_hydro.70.
DR Pfam; PF01473; CM binding.1; 12.
DR Pfam; PF02324; Glyco_hydro.70; 1.
KW Glycosyltransferase; Transferase.
SQ SEQUENCE 2057 AA; 228987 MW; 62BCE938SD9A11BE CRC64;

Query Match 57.5%; Score 65; DB 2; Length 2057;
Best Local Similarity 66.7%; Pred. No. 0.43;
Matches 12; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 4 HLSTLEAWSGNDNDYVK 21
Db 668 HVSILEDWSDNDADRYK 685

RESULT 15
08G9Q2
ID 08G9Q2 PRELIMINARY; PRT; 2835 AA.
AC 08G9Q2;
DT 01-MAR-2003 (TRENBLREL. 23, Created)
DT 01-MAR-2003 (TRENBLREL. 23, Last sequence update)
DT 01-MAR-2003 (TRENBLREL. 23, Last annotation update)
DE Dextranucrase (EC 2.4.1.5) (Fragment).
GN DSRE..
OS Leuconostoc mesenteroides.
OC Bacteria; Firmicutes; Lactobacillales; Leuconostoc.
OX NCBI_TaxID=1245;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=22231661; PubMed=12270834;
RA Bozonnet S., Dolis-Lafargue M., Fabre E., Pizzut S., Remaud-Simeon M.,
RA Mondan P., Willemot R.M.;
RT "Molecular characterization of DSR-E, an alpha-1,2 linkage
RT synthesizing dextranucrase with two catalytic domains.",
RL J. Bacteriol. 184:5753-5761(2002).
DR EMBL; AJ430204; CAD22883.1.
DR Transferase; Glycosyltransferase.
FT NON TER
SQ SEQUENCE 2835 AA; 313264 MW; D03262CD735399D CRC64;

Query Match 56.6%; Score 64; DB 2; Length 2835;
Best Local Similarity 55.0%; Pred. No. 0.9;
Matches 11; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 1 AIDHLSLEAWSGNDNDYVK 20
Db 557 ANNHISILEDWSDNDADRYK 576

Search completed: November 13, 2003, 09:44:06
Job time : 33.6493 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 13, 2003, 09:31:40 ; Search time 13.6351 Seconds
(without alignments)

65.165 Million cell updates/sec

Title: US-09-290-049a-13

Perfect score: 113

Sequence: 1 AIDHSLTEAWSGNDNDYKQ 21

Scoring table: BLOSUM62

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : Issued Patents AA:*

1: /cgn2_6/prodata/2/1aa/5A_COMB.pep:*\n2: /cgn2_6/prodata/2/1aa/5B_COMB.pep:*\n3: /cgn2_6/prodata/2/1aa/6A_COMB.pep:*\n4: /cgn2_6/prodata/2/1aa/6B_COMB.pep:*\n5: /cgn2_6/prodata/2/1aa/6CTUS_COMB.pep:*\n6: /cgn2_6/prodata/2/1aa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	75	66.4	545	US-09-604-957-4	Sequence 4, Appli
2	75	66.4	1430	US-09-009-172-2	Sequence 2, Appli
3	75	66.4	1430	US-09-210-361-6	Sequence 6, Appli
4	75	66.4	1430	US-09-740-274-6	Sequence 6, Appli
5	73	64.6	1475	US-09-007-999-2	Sequence 2, Appli
6	73	64.6	1475	US-09-210-361-2	Sequence 2, Appli
7	73	64.6	1475	US-09-740-274-2	Sequence 2, Appli
8	73	64.6	1577	US-08-793-824-2	Sequence 2, Appli
9	71	62.8	1375	US-09-210-361-4	Sequence 2, Appli
10	71	62.8	1375	US-09-740-274-4	Sequence 4, Appli
11	65	57.5	584	US-09-604-957-6	Sequence 6, Appli
12	65	57.5	2057	US-09-499-203-2	Sequence 2, Appli
13	58	51.3	523	US-09-604-957-5	Sequence 5, Appli
14	48	42.5	661	US-09-198-452A-36	Sequence 36, Appli
15	45	39.8	2366	US-08-480-604A-10	Sequence 10, Appli
16	45	39.8	2366	US-08-405-496A-10	Sequence 10, Appli
17	45	39.8	2366	US-08-915-136-10	Sequence 10, Appli
18	45	39.8	2366	US-08-957-310-10	Sequence 10, Appli
19	45	39.8	2366	US-10-011-366-10	Sequence 10, Appli
20	44	38.9	270	US-08-852-743-5	Sequence 5, Appli
21	44	38.9	270	US-09-185-370-5	Sequence 5, Appli
22	44	38.9	487	US-08-712-709-8	Sequence 8, Appli
23	44	38.9	487	US-09-111-444-8	Sequence 8, Appli
24	44	38.9	487	US-09-541-228-8	Sequence 8, Appli
25	44	38.9	639	US-08-557-309B-37	Sequence 37, Appli
26	44	38.9	639	US-08-834-306-37	Sequence 37, Appli
27	44	38.9	639	US-08-993-674A-37	Sequence 37, Appli

28	44	38.9	639	US-09-256-976-37	Sequence 37, Appli
29	43	38.1	225	US-09-107-532A-4105	Sequence 4105, Ap
30	43	38.1	456	US-09-172-841-51	Sequence 51, Appli
31	43	38.1	456	US-08-951-621-51	Sequence 51, Appli
32	43	38.1	535	US-09-604-957-7	Sequence 7, Appli
33	43	38.1	1278	US-09-604-957-3	Sequence 3, Appli
34	43	37.2	487	US-09-685-462-8	Sequence 8, Appli
35	42	37.2	1205	US-09-330-330-1	Sequence 1, Appli
36	41.5	36.7	679	US-09-302-620B-83	Sequence 83, Appli
37	41.5	36.7	679	US-09-302-620B-84	Sequence 84, Appli
38	41.5	36.7	679	US-09-912-161-4	Sequence 4, Appli
39	41.5	36.7	679	US-09-912-161-6	Sequence 6, Appli
40	41	36.3	132	US-09-073-297-16	Sequence 16, Appli
41	41	36.3	596	US-09-752-165-2	Sequence 2, Appli
42	41	36.3	833	US-09-514-302-3	Sequence 3, Appli
43	41	36.3	1938	US-09-514-302-2	Sequence 2, Appli
44	40	35.4	43	US-08-149-839B-8	Sequence 8, Appli
45	40	35.4	43	US-08-451-568-8	Sequence 8, Appli

ALIGNMENTS

```
RESULT 1
US-09-604-957-4 ; Application US/09604957
; Patent No. 6486314
; GENERAL INFORMATION:
; APPLICANT: VAN GEEL-SCHUTTEN, GERRITDINA HENDRIKA
; APPLICANT: DICHUIZEN, LOBBERT
; APPLICANT: RAHMOUI, HAKIM
; APPLICANT: LEER, ROBERT-JAN
; TITLE OF INVENTION: NOVEL GLUCOSYLTRANSFERASE AND NOVEL GLUCAN
; FILE REFERENCE: BO 43388
; CURRENT APPLICATION NUMBER: US/09/604,957
; PRIOR FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: 00201871.1
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 545
; TYPE: PRT
; ORGANISM: Streptococcus mutans
US-09-604-957-4

Query Match      66.4% ; Score 75 ; DB 4 ; Length 545 ;
Best Local Similarity 75.0% ; Pred. No. 0.00053 ;
Matches 15 ; Conservative 1 ; Mismatches 4 ; Indels 0 ; Gaps 0 ;
```

QY 1 AIDHSLTEAWSGNDNDYK 20
DB 75 AIDHSLTEAWSGNDNDYK 94

RESULT 2
US-09-008-172-2 ; Application US/09008172
; Patent No. 6127602
; GENERAL INFORMATION:
; APPLICANT: NICHOLS, Scott E.
; TITLE OF INVENTION: Substitutes for Modified Starch and
; FILE REFERENCE: 0358D
; CURRENT APPLICATION NUMBER: US/09/008,172
; PRIOR FILING DATE: 1998-01-16
; EARLIER APPLICATION NUMBER: 08/482,711
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 1430
; TYPE: PRT

ORGANISM: Streptococcus mutans
US-09-008-172-2

Query Match 66.4%; Score 75; DB 3; Length 1430;
Best Local Similarity 75.0%; Pred. No. 0.0017;
Matches 15; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 AIDHLSILEAMSGNDNDYK 20
|:|||||:|
Db 495 AINHLSILEAMSDNDPQYK 514

RESULT 3
US-09-210-361-6

Sequence 6, Application US/09210361
Patent No. 6284479
GENERAL INFORMATION:
APPLICANT: Nichols, Scott E.
TITLE OF INVENTION: Substitutes for Modified Starches and
FILE REFERENCE: 0357CR

CURRENT APPLICATION NUMBER: US/09/210,361
EARLIER FILING DATE: 1998-12-11
EARLIER APPLICATION NUMBER: 09/007,999
EARLIER FILING DATE: 1998-01-16
EARLIER APPLICATION NUMBER: 08/478,704
EARLIER FILING DATE: 1995-06-07
EARLIER APPLICATION NUMBER: 09/009,620
EARLIER FILING DATE: 1998-01-20
EARLIER APPLICATION NUMBER: 08/485,243
EARLIER FILING DATE: 1995-06-07
EARLIER APPLICATION NUMBER: 09/008,172
EARLIER FILING DATE: 1998-01-16
EARLIER APPLICATION NUMBER: 08/482,711
EARLIER FILING DATE: 1995-06-07
NUMBER OF SEQ ID NOS: 6
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 6
LENGTH: 1430
TYPE: PRT
ORGANISM: streptococcus mutans
US-09-210-361-6

Query Match 66.4%; Score 75; DB 3; Length 1430;
Best Local Similarity 75.0%; Pred. No. 0.0017;
Matches 15; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 AIDHLSILEAMSGNDNDYK 20
|:|||||:|
Db 495 AINHLSILEAMSDNDPQYK 514

RESULT 4

US-09-740-274-6

Sequence 6, Application US/09740274
Patent No. 6465203
GENERAL INFORMATION:
APPLICANT: Nichols, Scott E.
TITLE OF INVENTION: Glucan-containing Compositions and Paper
FILE REFERENCE: 0357CR
CURRENT APPLICATION NUMBER: US/09/740,274
CURRENT FILING DATE: 2000-12-19
EARLIER APPLICATION NUMBER: 09/210,361
EARLIER FILING DATE: 1998-12-11
EARLIER APPLICATION NUMBER: 09/007,999
EARLIER FILING DATE: 1998-01-16
EARLIER APPLICATION NUMBER: 08/478,704
EARLIER FILING DATE: 1995-06-07
EARLIER APPLICATION NUMBER: 09/009,620
EARLIER FILING DATE: 1998-01-20
EARLIER APPLICATION NUMBER: 08/485,243
EARLIER FILING DATE: 1995-06-07
EARLIER APPLICATION NUMBER: 09/008,172

PRIOR FILING DATE: 1998-01-16
PRIOR APPLICATION NUMBER: 08/482,711
PRIOR FILING DATE: 1995-06-07
NUMBER OF SEQ ID NOS: 6
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 6
LENGTH: 1430
TYPE: PRT
ORGANISM: streptococcus mutans
US-09-740-274-6

Query Match 66.4%; Score 75; DB 4; Length 1430;
Best Local Similarity 75.0%; Pred. No. 0.0017;
Matches 15; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 AIDHLSILEAMSGNDNDYK 20
|:|||||:|
Db 495 AINHLSILEAMSDNDPQYK 514

RESULT 5

US-09-007-999-2

Sequence 2, Application US/09007999
Patent No. 6087559
GENERAL INFORMATION:
APPLICANT: Nichols, Scott E.
TITLE OF INVENTION: Substitutes for Modified Starch and
FILE REFERENCE: 0356D
CURRENT APPLICATION NUMBER: US/09/007,999
CURRENT FILING DATE: 1998-01-16
EARLIER APPLICATION NUMBER: 08/478,704
EARLIER FILING DATE: 1995-06-07
NUMBER OF SEQ ID NOS: 2
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 2
LENGTH: 1475
TYPE: PRT
ORGANISM: Streptococcus mutans
US-09-007-999-2

Query Match 64.6%; Score 73; DB 3; Length 1475;
Best Local Similarity 73.7%; Pred. No. 0.0036;
Matches 14; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 AIDHLSILEAMSGNDNDYV 19
|:|||||:|
Db 481 ANDHLSILEAMSDNDTPYL 499

RESULT 6

US-09-210-361-2

Sequence 2, Application US/09210361
Patent No. 6284479
GENERAL INFORMATION:
APPLICANT: Nichols, Scott E.
TITLE OF INVENTION: Substitutes for Modified Starches and
FILE REFERENCE: 0357CR
CURRENT APPLICATION NUMBER: US/09/210,361
CURRENT FILING DATE: 1998-12-11
EARLIER APPLICATION NUMBER: 09/007,999
EARLIER FILING DATE: 1998-01-16
EARLIER APPLICATION NUMBER: 08/478,704
EARLIER FILING DATE: 1995-06-07
EARLIER APPLICATION NUMBER: 09/009,620
EARLIER FILING DATE: 1998-01-20
EARLIER APPLICATION NUMBER: 08/485,243
EARLIER FILING DATE: 1995-06-07
EARLIER APPLICATION NUMBER: 09/008,172
EARLIER FILING DATE: 1998-01-16
EARLIER APPLICATION NUMBER: 08/482,711

```

; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 1475
; TYPE: PRF
; ORGANISM: Streptococcus mutans
US-09-210-361-2

Query Match      64.6%; Score 73; DB 3; Length 1475;
Best Local Similarity 73.7%; Pred. No. 0.0036;
Matches 14; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 AIDHLSILAWSGNDNDYV 19
Db 481 ANDHLSILAWSGNDNDYV 499

RESULT 7
US-09-740-274-2
; Sequence 2, Application US/09740274
; Patent No. 6465203
; GENERAL INFORMATION:
; APPLICANT: Nichols, Scott E.
; TITLE OF INVENTION: Glucan-containing Compositions and Paper
; FILE REFERENCE: 0357CRD
; CURRENT APPLICATION NUMBER: US/09/740,274
; PRIOR FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: 09/210,361
; PRIOR FILING DATE: 1998-12-11
; PRIOR APPLICATION NUMBER: 09/007,999
; PRIOR FILING DATE: 1998-01-16
; PRIOR APPLICATION NUMBER: 08/478,704
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: 09/009,620
; PRIOR FILING DATE: 1998-01-20
; PRIOR APPLICATION NUMBER: 08/485,243
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: 09/008,172
; PRIOR FILING DATE: 1998-01-16
; PRIOR APPLICATION NUMBER: 08/482,711
; PRIOR FILING DATE: 1995-06-07
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 1475
; TYPE: PRF
; ORGANISM: Streptococcus mutans
US-09-740-274-2

Query Match      64.6%; Score 73; DB 4; Length 1475;
Best Local Similarity 73.7%; Pred. No. 0.0036;
Matches 14; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 AIDHLSILAWSGNDNDYV 19
Db 481 ANDHLSILAWSGNDNDYV 499

RESULT 8
US-08-793-824-2
; Sequence 2, Application US/08793824
; Patent No. 5981838
; GENERAL INFORMATION:
; APPLICANT: Simpson, Christine Lynn
; APPLICANT: Giffard, Philip Morrison
; APPLICANT: Jacques, Nicholas Anthony
; TITLE OF INVENTION: Genetic Manipulation of Plants to
; TITLE OF INVENTION: Increase Stored Carbohydrates
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Griffith Hack & Co
; STREET: Level 8, 168 Walker Street
; CITY: No. 5981838th Sydney
```

```

; STATE: New South Wales
; COUNTRY: Australia
; ZIP: 2060
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/793,824
; FILING DATE:
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: AU PM7643
; FILING DATE: 24-AUG-1994
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 61 2 957 5944
; TELEFAX: 61 2 957 6288
; TELEX: 26547
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1577 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: not relevant
; MOLECULE TYPE: protein
; ORIGINAL SOURCE:
; ORGANISM: Streptococcus salivarius
US-08-793-824-2

Query Match      64.6%; Score 73; DB 2; Length 1577;
Best Local Similarity 75.0%; Pred. No. 0.0039;
Matches 15; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 AIDHLSILAWSGNDNDYK 20
Db 588 AIAHLSILAWSGNDNDYK 607

RESULT 9
US-09-210-361-4
; Sequence 4, Application US/09210361
; Patent No. 6284479
; GENERAL INFORMATION:
; APPLICANT: Nichols, Scott E.
; TITLE OF INVENTION: Substitutes for Modified Starches and
; TITLE OF INVENTION: Latexes in Paper Manufacture
; FILE REFERENCE: 0357CR
; CURRENT APPLICATION NUMBER: US/09/210,361
; CURRENT FILING DATE: 1998-12-11
; EARLIER APPLICATION NUMBER: 09/007,999
; EARLIER FILING DATE: 1998-01-16
; EARLIER APPLICATION NUMBER: 08/478,704
; EARLIER FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 09/009,620
; EARLIER FILING DATE: 1998-01-20
; EARLIER APPLICATION NUMBER: 08/485,243
; EARLIER FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 09/008,172
; EARLIER FILING DATE: 1998-01-16
; EARLIER APPLICATION NUMBER: 08/482,711
; EARLIER FILING DATE: 1995-06-07
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 1375
; TYPE: PRF
; ORGANISM: streptococcus mutans
US-09-210-361-4

Query Match      62.8%; Score 71; DB 3; Length 1375;
Best Local Similarity 73.7%; Pred. No. 0.007;
Matches 14; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
```

```
QY      1 AIDHLSILEAWSGNDNDYV 19
          | | | | | | | | | :
Db      507 ANDHLSILEAWSYNDDPYL 525
```

```

RESULT 10 274-4
US-09-740-274-4
: Sequence 4, Application US/09740274
: Patent No. 6465203
: GENERAL INFORMATION:
: APPLICANT: Nicholas, Scott E.
: TITLE OF INVENTION: Glucan-containing Compositions and Paper
: FILE REFERENCE: 0357CRD
: CURRENT APPLICATION NUMBER: US/09/740,274
: CURRENT FILING DATE: 2000-12-19
: PRIOR APPLICATION NUMBER: 03/210,361
: PRIOR FILING DATE: 1998-12-11
: PRIOR APPLICATION NUMBER: 09/007,999
: PRIOR FILING DATE: 1998-01-16
: PRIOR APPLICATION NUMBER: 08/478,704
: PRIOR FILING DATE: 1995-06-07
: PRIOR APPLICATION NUMBER: 03/009,620
: PRIOR FILING DATE: 1998-01-20
: PRIOR APPLICATION NUMBER: 08/485,243
: PRIOR FILING DATE: 1995-06-07
: PRIOR APPLICATION NUMBER: 09/008,172
: PRIOR FILING DATE: 1998-01-16
: PRIOR APPLICATION NUMBER: 08/482,711
: PRIOR FILING DATE: 1995-06-07
: NUMBER OF SEQ ID NOS: 6
: SOFTWARE: FastSeq for Windows Version 3.0
: SEQ ID NO 4
: LENGTH: 1375
: TYPE: PRT
: ORGANISM: streptococcus mutans
US-09-740-274-4

```

Query Match	62.8%	Score 71	DB 4	Length 1375
Best Local Similarity	73.7%	Pred. No.	0.007	
Matches 14	Conservative 1	Mismatches 4	Indels 0	Gaps 0

```
QY      1 AIDHLSILEAWSGNDNDYV 19
          |||||
Db      507 ANDHLSILEAWSYNDTPYL 525
```

```

RESULT 11.957-6
US-09-604-957-6
: Sequence 6, Application US/09604957
: Patent No. 6486314
: GENERAL INFORMATION:
: APPLICANT: VAN GEEL-SCHUTTEN, GERRITJINA HENDRIKA
: APPLICANT: DIKHUIZEN, LUBBERT
: APPLICANT: RAHAOUI, HAKIM
: APPLICANT: LEER, ROBERT-JAN
: TITLE OF INVENTION: NOVEL GLUCOSYLTRANSFERASE AND NOVEL GLUCAN
: FILE REFERENCE: BO 4338
: CURRENT APPLICATION NUMBER: US/09/604,957
: CURRENT FILING DATE: 2000-06-28
: PRIOR APPLICATION NUMBER: 00201871.1
: PRIOR FILING DATE: 2000-05-25
: NUMBER OF SEQ. ID NOS.: 17
: SOFTWARE: PatentIn Ver. 2.1
: SEQ ID NO. 6
: LENGTH: 584
: TYPE: PRT
: ORGANISM: Leuconostoc mesenteroides
: US-09-604-957-6

```

Query Match	57.5%	Score 65;	DB 4;	Length 584;
Best Local Similarity	66.7%	Pred. No. 0.023;		
Matches 12;	Conservative 1;	Mismatches 5;	Indels 0;	Gaps 0;

```
QY      4 HSLLEAWSGNDNDYKQ 21
          |||||:|||||
Db      78 HSLLEDWNGKDPQYVWQ 95
```

```

RESULT 12
US-09-499-203-2
; Sequence 2, Application US/09499203
; Patent No. 6570065
; GENERAL INFORMATION:
; APPLICANT: KOSSMANN, Jens
; APPLICANT: WELSH, Thomas
; APPLICANT: QUANZ, Martin
; APPLICANT: KNUTH, Karola
; TITLE OF INVENTION: Nucleic Acid Molecules Encoding Alternansucrase
; FILE REFERENCE: 147-196P
; CURRENT APPLICATION NUMBER: US/09/499,203
; CURRENT FILING DATE: 2000-02-08
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 2057
; TYPE: PRT
; ORGANISM: Leuconostoc mesenteroides
US-09-499-203-2

```

Query Match	57.5%	Score 65	DB 4	Length 2057
Best Local Similarity	66.7%	Pred. No.	0.1	
Matches 12	Conservative 1	Mismatches 5	Indels 0	Gaps 0

QY 4 HLSILEAWSGNDNDYKQ 21
||| | : | | |
Db 668 HLSILEDWNGKDPQYVQ 685

RESULT 13
US-09-604-957-5
Sequence 5, Application US/09604957
Patent No. 6486514
GENERAL INFORMATION:
APPLICANT: VAN GEEL-SCHUTTEN, GERRITJINA HENDRIKA
APPLICANT: DIJKHUIZEN, LUBBERT
APPLICANT: RAHAOUTI, HAKIM
APPLICANT: LEER, ROBERT-JAN
TITLE OF INVENTION: NOVEL GLUCOSYLTRANSFERASE AND NOVEL GLUCAN
FILE REFERENCE: BO 43388
CURRENT APPLICATION NUMBER: US/09/604,957
CURRENT FILING DATE: 2000-06-28
PRIOR APPLICATION NUMBER: 00201871.1
PRIOR FILING DATE: 2000-05-25
NUMBER OF SEQ ID NOS: 17
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 5
LENGTH: 523
TYPE: prt
ORGANISM: Leuconostoc mesenteroides
US-09-604-957-5

Query Match	51.3%	Score	58	DB	4	Length	523
Best Local Similarity	68.4%	Pred. NO.	0.27				
Matches	13	Conservative	0	Mismatches	6	Indels	0
						Gaps	0

QY 1 AIDHLSILEAWSGNDNDYV 19
| | | | | | | | | |
Db 75 ANQHSILEEDWSHNDPLYV 93

RESULT 14
US-09-198-452A-36
; Sequence 36, Application US/09198452A
; Patent No. 6559294
; GENERAL INFORMATION:

APPLICANT: Grifflais, R.
TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragments
TITLE OF INVENTION: thereof and uses thereof, in particular for the diagnosis, prevention
TITLE OF INVENTION: and treatment of infection
FILE REFERENCE: 9710-003-999
CURRENT APPLICATION NUMBER: US/09/198,452A
CURRENT FILING DATE: 1998-11-24
NUMBER OF SEQ ID NOS: 6849
SEQ ID NO: 36
LENGTH: 661
TYPE: PRT
ORGANISM: Chlamydia pneumoniae
US-09-198-452A-36

Query Match 42.5%; Score 48; DB 4; Length 661;
Best Local Similarity 57.1%; Pred. No. 15;
Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Oy 3 DHLSLEAMSGNDN 16
Db 572 DHMNRWEMIGNDN 585

RESULT 15
US-08-480-604A-10
Sequence 10, Application US/08480604A
Patent No. 5736139
GENERAL INFORMATION:
APPLICANT: KINK, JOHN A.
APPLICANT: THALLEY, BRUCE S.
APPLICANT: PADHYE, NISHA V.
APPLICANT: PIRCA, JOSEPH R.
APPLICANT: STAFFORD, DOUGLAS C.
TITLE OF INVENTION: VACCINE AND ANTITOXIN FOR TREATMENT AND
NUMBER OF SEQUENCES: 32
CORRESPONDENCE ADDRESS:
ADDRESSEE: MEDLEN & CARROLL, LLP
STREET: 220 MONTGOMERY STREET, SUITE 2200
CITY: SAN FRANCISCO
STATE: CALIFORNIA
COUNTRY: UNITED STATES OF AMERICA
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/480,604A
FILING DATE: 07-JUN-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/422,711
FILING DATE: 14-APR-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/405,496
FILING DATE: 16-MAR-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/329,154
FILING DATE: 25-OCT-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/161,907
FILING DATE: 02-DEC-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/985,321
FILING DATE: 04-DEC-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/429,791
FILING DATE: 31-OCT-1989
ATTORNEY/AGENT INFORMATION:
NAME: INGOLIA, DIANE E.
REGISTRATION NUMBER: 40,027

REFERENCE/DOCKET NUMBER: OPHD-01763
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 705-8410
TELEFAX: (415) 397-8338
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 2366 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-480-604A-10

Query Match 39.8%; Score 45; DB 1; Length 2366;
Best Local Similarity 44.4%; Pred. No. 2e+02;
Matches 8; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

Oy 2 IDHLSLEAMSGNDNDYV 19
Db 1733 INDLSIRYWSNDGNDPI 1750

Search completed: November 13, 2003, 09:48:01
Job time : 14.6351 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 13, 2003, 09:45:40 / Search time 27.4692 Seconds
(without alignments)
139.565 Million cell updates/sec

Title: US-09-290-049A-13

Perfect score: 113

Sequence: 1 AIDHSLTEAWSGNDNDYVKQ 21

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 66188 seqs, 182559486 residues

Total number of hits satisfying chosen parameters: 66188

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA:*

1: /cgn2_6/ptodata/2/pubppaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/2/pubppaa/PC1_NEW_PUB.pep:*
3: /cgn2_6/ptodata/2/pubppaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/2/pubppaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/2/pubppaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/2/pubppaa/CT05_PUBCOMB.pep:*
7: /cgn2_6/ptodata/2/pubppaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/2/pubppaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/2/pubppaa/US09_PUBCOMB.pep:*
10: /cgn2_6/ptodata/2/pubppaa/US09_PUBCOMB.pep:*
11: /cgn2_6/ptodata/2/pubppaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/2/pubppaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/2/pubppaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/2/pubppaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/2/pubppaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/2/pubppaa/US10_NEW_PUB.pep:*
17: /cgn2_6/ptodata/2/pubppaa/US60_NEW_PUB.pep:*
18: /cgn2_6/ptodata/2/pubppaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	75	66.4	545	10	US-09-995-749A-10
2	75	66.4	1430	9	US-09-740-274-6
3	73	64.6	1475	9	US-09-740-274-2
4	71	62.8	1375	9	US-09-740-274-4
5	65	57.5	584	10	US-09-995-749A-12
6	58	51.3	522	10	US-09-995-749A-11
7	48.5	42.9	427	11	US-09-813-153-131
8	45	39.8	58	9	US-09-764-887-163
9	45	39.8	58	15	US-10-073-961-163
10	45	39.8	2366	15	US-10-011-366-10
11	44	38.9	207	12	US-10-225-810-28
12	44	38.9	448	12	US-10-225-810-44
13	44	38.9	487	9	US-09-810-808-8
14	44	38.9	554	11	US-09-813-409-28
15	43	38.1	272	12	US-10-291-253A-11

16	43	38.1	293	12	US-10-291-253A-10	Sequence 10, Appl
17	43	38.1	535	10	US-09-995-749A-13	Sequence 13, Appl
18	43	38.1	659	12	US-10-214-529-3	Sequence 3, Appl
19	43	38.1	1044	9	US-09-822-268A-5	Sequence 5, Appl
20	43	38.1	1170	9	US-09-822-268A-4	Sequence 4, Appl
21	43	38.1	1172	9	US-09-822-268A-2	Sequence 2, Appl
22	43	38.1	1237	12	US-10-154-419-82	Sequence 82, Appl
23	43	38.1	1237	12	US-10-146-733-77	Sequence 77, Appl
24	43	38.1	1237	14	US-10-024-623-32	Sequence 32, Appl
25	43	38.1	1781	10	US-09-995-749A-2	Sequence 2, Appl
26	42	37.2	116	15	US-10-078-770-196	Sequence 196, Appl
27	42	37.2	160	9	US-09-860-232A-17	Sequence 17, Appl
28	42	37.2	235	16	US-10-080-170-550	Sequence 550, Appl
29	42	37.2	240	10	US-09-738-626-6348	Sequence 6348, Appl
30	42	37.2	388	10	US-09-738-626-5242	Sequence 5242, Appl
31	42	37.2	433	14	US-10-078-829-172	Sequence 172, Appl
32	42	37.2	500	15	US-10-156-761-8558	Sequence 8558, Appl
33	42	37.2	649	15	US-10-078-770-192	Sequence 4, Appl
34	42	37.2	695	15	US-10-139-496-4	Sequence 4579, Appl
35	42	37.2	794	10	US-09-738-626-4579	Sequence 3018, Appl
36	42	37.2	867	15	US-10-128-714-3018	Sequence 8018, Appl
37	42	37.2	974	15	US-10-128-714-8018	Sequence 7, Appl
38	42	37.2	1201	12	US-10-214-529-7	Sequence 655, Appl
39	42	37.2	1915	12	US-10-238-075-655	Sequence 2, Appl
40	41.5	36.7	648	15	US-10-272-017A-3	Sequence 3, Appl
41	41.5	36.7	648	15	US-10-272-017A-2	Sequence 5, Appl
42	41.5	36.7	648	15	US-10-272-017A-5	Sequence 6, Appl
43	41.5	36.7	648	15	US-10-272-017A-6	Sequence 11, Appl
44	41.5	36.7	679	9	US-09-911-781-11	Sequence 83, Appl
45	41.5	36.7	679	11	US-09-976-800-83	

ALIGNMENTS

RESULT 1
US-09-995-749A-10
Sequence 10, Application US/0995749A
Patent No. US20020155568A1

GENERAL INFORMATION: GEBEL-SCHUTTEN, GERRITDINA HENDRIKA
APPLICANT: VAN GEBEL-SCHUTTEN, GERRITDINA HENDRIKA
APPLICANT: DIKHUIZEN, LOUBERT
APPLICANT: RAHMOUI, HAKIM
APPLICANT: LEEER, ROBERT-JAN
TITLE OF INVENTION: NOVEL GLUCOSYLTRANSFERASIS
FILE REFERENCE: BO43388-CIP
CURRENT FILING DATE: 2001-11-29
PRIOR APPLICATION NUMBER: US/09/995,749A
PRIOR FILING DATE: 2000-06-28
PRIOR APPLICATION NUMBER: EPO 00201871.1
PRIORITY FILING DATE: 2000-05-25
NUMBER OF SEQ ID NOS: 19
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 10
LENGTH: 545
TYPE: PRT
ORGANISM: Streptococcus mutans
US-09-995-749A-10

Query Match 66.4%; Score 75; DB 10; Length 545;
Best Local Similarity 75.0%; Pred. No. 0.0034;
Matches 15; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 AIDHSLTEAWSGNDNDYVK 20
DB 75 AIDHSLTEAWSGNDNDYVK 94

RESULT 2
US-09-740-274-6
Sequence 6, Application US/09740274
Patent No. US20020031826A1

```
; GENERAL INFORMATION:
; APPLICANT: Nichols, Scott E.
; TITLE OF INVENTION: Glucan-containing Compositions and Paper
; FILE REFERENCE: 0357CRD
; CURRENT APPLICATION NUMBER: US/09/740,274
; CURRENT FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: 09/210,361
; PRIOR FILING DATE: 1998-12-11
; PRIOR APPLICATION NUMBER: 09/007,999
; PRIOR FILING DATE: 1998-01-16
; PRIOR APPLICATION NUMBER: 08/478,704
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: 09/009,620
; PRIOR FILING DATE: 1998-01-20
; PRIOR APPLICATION NUMBER: 08/485,243
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: 08/478,704
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: 09/008,172
; PRIOR FILING DATE: 1998-01-16
; PRIOR APPLICATION NUMBER: 08/482,711
; PRIOR FILING DATE: 1995-06-07
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 6
; LENGTH: 1430
; TYPE: PRT
; ORGANISM: Streptococcus mutans
; US-09-740-274-6

Query Match          66.4%; Score 75; DB 9; Length 1430;
Best Local Similarity 75.0%; Pred. No. 0.0095;
Matches 15; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY      1 AIDHSLTEAMSGNDNDYK 20
Db      495 AINHLSTLEAMSGNDNDPQYNK 514

RESULT 3
US-09-740-274-2
; Sequence 2, Application US/09740274
; Patent No. US20020031826A1
; GENERAL INFORMATION:
; APPLICANT: Nichols, Scott E.
; TITLE OF INVENTION: Glucan-containing Compositions and Paper
; FILE REFERENCE: 0357CRD
; CURRENT APPLICATION NUMBER: US/09/740,274
; CURRENT FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: 09/210,361
; PRIOR FILING DATE: 1998-12-11
; PRIOR APPLICATION NUMBER: 09/007,999
; PRIOR FILING DATE: 1998-01-16
; PRIOR APPLICATION NUMBER: 08/478,704
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: 09/009,620
; PRIOR FILING DATE: 1998-01-20
; PRIOR APPLICATION NUMBER: 08/485,243
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: 09/008,172
; PRIOR FILING DATE: 1998-01-16
; PRIOR APPLICATION NUMBER: 08/482,711
; PRIOR FILING DATE: 1995-06-07
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 1475
; TYPE: PRT
; ORGANISM: Streptococcus mutans
; US-09-740-274-2

Query Match          64.6%; Score 73; DB 9; Length 1475;
Best Local Similarity 73.7%; Pred. No. 0.02;
Matches 14; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
```

```
QY      1 AIDHSLTEAMSGNDNDYV 19
Db      481 ANDHSLTEAMSGNDNDTPYL 499

RESULT 4
US-09-740-274-4
; Sequence 4, Application US/09740274
; Patent No. US20020031826A1
; GENERAL INFORMATION:
; APPLICANT: Nichols, Scott E.
; TITLE OF INVENTION: Glucan-containing Compositions and Paper
; FILE REFERENCE: 0357CRD
; CURRENT APPLICATION NUMBER: US/09/740,274
; CURRENT FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: 09/210,361
; PRIOR FILING DATE: 1998-12-11
; PRIOR APPLICATION NUMBER: 09/007,999
; PRIOR FILING DATE: 1998-01-16
; PRIOR APPLICATION NUMBER: 08/478,704
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: 09/009,620
; PRIOR FILING DATE: 1998-01-20
; PRIOR APPLICATION NUMBER: 08/485,243
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: 09/008,172
; PRIOR FILING DATE: 1998-01-16
; PRIOR APPLICATION NUMBER: 08/482,711
; PRIOR FILING DATE: 1995-06-07
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 1375
; TYPE: PRT
; ORGANISM: Streptococcus mutans
; US-09-740-274-4

Query Match          62.8%; Score 71; DB 9; Length 1375;
Best Local Similarity 73.7%; Pred. No. 0.037;
Matches 14; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY      1 AIDHSLTEAMSGNDNDYV 19
Db      507 ANDHSLTEAMSGNDNDTPYL 525

RESULT 5
US-09-995-749A-12
; Sequence 12, Application US/0995749A
; Patent No. US2002015568A1
; GENERAL INFORMATION:
; APPLICANT: VAN GEEL-SCHUTTEN, GERRITDINA HENDRIKA
; APPLICANT: DIJKHUIZEN, LUBBERT
; APPLICANT: RAHOUI, HAKIM
; APPLICANT: LEER, ROBERT-JAN
; TITLE OF INVENTION: NOVEL GLUCOSYLTRANSFERASES
; FILE REFERENCE: B043388-CIP
; CURRENT APPLICATION NUMBER: US/09/995,749A
; CURRENT FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: 09/604,957
; PRIOR FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: EPO 00201871.1
; PRIOR FILING DATE: 2000-05-25
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 12
; LENGTH: 584
; TYPE: PRT
; ORGANISM: Leuconostoc mesenteroides
; US-09-995-749A-12

Query Match          57.5%; Score 65; DB 10; Length 584;
Best Local Similarity 66.7%; Pred. No. 0.12;
```

```
Matches 12; Conservative 1; Mismatches 5; Indels 0; Gaps 0;
QY 4 HSLILEAMSGNDNDYKQ 21
    |||||:|:|
Db 78 HSLILEDMWNGKDPQYVQ 95

RESULT 6
US-09-995-749A-11
; Sequence 11, Application US/09995749A
; Patent No. US2002015568A1
; GENERAL INFORMATION:
; APPLICANT: VAN GEEL-SCHUTTEN, GERRITDINA, HENDRIKA
; APPLICANT: DUTKHUIZEN, LUBBERT
; APPLICANT: RAHAOUI, HAKIM
; APPLICANT: LEER, ROBERT-JAN
; TITLE OF INVENTION: NOVEL GLYCOSYLTRANSFERASES
; FILE REFERENCE: B043388-CIP
; CURRENT APPLICATION NUMBER: US/09/995,749A
; CURRENT FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: 09/604,957
; PRIOR FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: EPO 00201871.1
; PRIOR FILING DATE: 2000-05-25
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 11
; LENGTH: 522
; TYPE: PRT
; ORGANISM: Leuconostoc mesenteroides
US-09-995-749A-11

Query Match 51.3%; Score 58; DB 10; Length 522;
Best Local Similarity 68.4%; Pred. No. 1.2;
Matches 13; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
QY 1 AIDHSLILEAMSGNDNDYV 19
    |||||:|:|
Db 75 ANQHSLILEDMWSHNDPLYV 93

RESULT 7
US-09-813-153-131
; Sequence 131, Application US/09813153
; Publication No. US20030045459A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 67 Human secreted proteins
; FILE REFERENCE: P2023
; CURRENT APPLICATION NUMBER: US/09/813,153
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US/09/363,044
; PRIOR FILING DATE: 1999-07-29
; PRIOR APPLICATION NUMBER: 60/073,160
; PRIOR FILING DATE: 1998-01-30
; PRIOR APPLICATION NUMBER: 60/073,159
; PRIOR FILING DATE: 1998-01-30
; PRIOR APPLICATION NUMBER: 60/073,165
; PRIOR FILING DATE: 1998-01-30
; PRIOR APPLICATION NUMBER: 60/073,164
; PRIOR FILING DATE: 1998-01-30
; PRIOR APPLICATION NUMBER: 60/073,167
; PRIOR FILING DATE: 1998-01-30
; PRIOR APPLICATION NUMBER: 60/073,162
; PRIOR FILING DATE: 1998-01-30
; PRIOR APPLICATION NUMBER: 60/073,161
; PRIOR FILING DATE: 1998-01-30
; PRIOR APPLICATION NUMBER: 60/073,170
; PRIOR FILING DATE: 1998-01-30
; NUMBER OF SEQ ID NOS: 298
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 131
; LENGTH: 427
```

```
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (427)
; OTHER INFORMATION: Xaa equals stop translation
US-09-813-153-131

Query Match 42.9%; Score 48.5; DB 11; Length 427;
Best Local Similarity 52.6%; Pred. No. 28;
Matches 10; Conservative 4; Mismatches 4; Indels 1; Gaps 1;
QY 3 DRLVLEBSWA-NDDPYLKR 21
    |||||:|:|
Db 269 DRLVLEBSWA-NDDPYLKR 286

RESULT 8
US-09-764-887-163
; Sequence 163, Application US/09764887
; Patent No. US20020042096A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: P413
; CURRENT APPLICATION NUMBER: US/09/764,887
; CURRENT FILING DATE: 2001-01-17
; PRIOR APPLICATION data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 658
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 163
; LENGTH: 58
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (48)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; LOCATION: (51)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-764-887-163

Query Match 39.8%; Score 45; DB 9; Length 58;
Best Local Similarity 56.2%; Pred. No. 11;
Matches 9; Conservative 1; Mismatches 6; Indels 0; Gaps 0;
QY 1 AIDHSLILEAMSGNDN 16
    |||||:|:|
Db 25 AAGSLILEGWSGEDH 40

RESULT 9
US-10-073-961-163
; Sequence 163, Application US/10073961
; Publication No. US20030077602A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: P413C1
; CURRENT APPLICATION NUMBER: US/10/073,961
; CURRENT FILING DATE: 2002-03-05
; PRIOR APPLICATION NUMBER: 09/764,887
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 60/179,065
; PRIOR FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: 60/180,628
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: 60/214,886
; PRIOR FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: 60/217,487
; PRIOR FILING DATE: 2000-07-11
; PRIOR APPLICATION NUMBER: 60/225,758
```

PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/220,963
PRIOR FILING DATE: 2000-07-26
PRIOR APPLICATION NUMBER: 60/217,496
PRIOR FILING DATE: 2000-07-11
PRIOR APPLICATION NUMBER: 60/225,447
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/218,290
PRIOR FILING DATE: 2000-07-14
PRIOR APPLICATION NUMBER: 60/225,757
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/226,868
PRIOR FILING DATE: 2000-08-22
PRIOR APPLICATION NUMBER: 60/216,647
PRIOR FILING DATE: 2000-07-07
PRIOR APPLICATION NUMBER: 60/225,270
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/251,869
PRIOR FILING DATE: 2000-12-08
PRIOR APPLICATION NUMBER: 60/235,834
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: 60/234,274
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: 60/234,223
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: 60/228,924
PRIOR FILING DATE: 2000-08-30
PRIOR APPLICATION NUMBER: 60/224,518
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/226,369
PRIOR FILING DATE: 2000-09-29
PRIOR APPLICATION NUMBER: 60/224,519
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/220,964
PRIOR FILING DATE: 2000-07-26
PRIOR APPLICATION NUMBER: 60/241,809
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/249,299
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/236,327
PRIOR FILING DATE: 2000-09-29
PRIOR APPLICATION NUMBER: 60/241,785
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/244,617
PRIOR FILING DATE: 2000-11-01
PRIOR APPLICATION NUMBER: 60/225,268
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/236,368
PRIOR FILING DATE: 2000-09-29
PRIOR APPLICATION NUMBER: 60/251,856
PRIOR FILING DATE: 2000-12-08
PRIOR APPLICATION NUMBER: 60/251,868
PRIOR FILING DATE: 2000-12-08
PRIOR APPLICATION NUMBER: 60/229,344
PRIOR FILING DATE: 2000-09-01
PRIOR APPLICATION NUMBER: 60/234,997
PRIOR FILING DATE: 2000-09-25
PRIOR APPLICATION NUMBER: 60/229,343
PRIOR FILING DATE: 2000-09-01
PRIOR APPLICATION NUMBER: 60/229,345
PRIOR FILING DATE: 2000-09-01
PRIOR APPLICATION NUMBER: 60/229,287
PRIOR FILING DATE: 2000-09-01
PRIOR APPLICATION NUMBER: 60/229,513
PRIOR FILING DATE: 2000-09-05
PRIOR APPLICATION NUMBER: 60/231,413
PRIOR FILING DATE: 2000-09-08
PRIOR APPLICATION NUMBER: 60/229,509
PRIOR FILING DATE: 2000-09-05

PRIOR APPLICATION NUMBER: 60/236,367
PRIOR FILING DATE: 2000-09-29
PRIOR APPLICATION NUMBER: 60/237,039
PRIOR FILING DATE: 2000-10-02
PRIOR APPLICATION NUMBER: 60/237,038
PRIOR FILING DATE: 2000-10-02
PRIOR APPLICATION NUMBER: 60/236,370
PRIOR FILING DATE: 2000-09-29
PRIOR APPLICATION NUMBER: 60/236,802
PRIOR FILING DATE: 2000-10-02
PRIOR APPLICATION NUMBER: 60/237,037
PRIOR FILING DATE: 2000-10-02
PRIOR APPLICATION NUMBER: 60/237,040
PRIOR FILING DATE: 2000-10-02
PRIOR APPLICATION NUMBER: 60/240,960
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/239,935
PRIOR FILING DATE: 2000-10-13
PRIOR APPLICATION NUMBER: 60/239,937
PRIOR FILING DATE: 2000-10-13
PRIOR APPLICATION NUMBER: 60/241,787
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/246,474
PRIOR FILING DATE: 2000-11-08
PRIOR APPLICATION NUMBER: 60/246,532
PRIOR FILING DATE: 2000-11-08
PRIOR APPLICATION NUMBER: 60/249,216
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/249,210
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/226,681
PRIOR FILING DATE: 2000-08-22
PRIOR APPLICATION NUMBER: 60/225,759
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/225,213
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/227,182
PRIOR FILING DATE: 2000-08-22
PRIOR APPLICATION NUMBER: 60/225,214
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/235,836
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: 60/230,438
PRIOR FILING DATE: 2000-09-06
PRIOR APPLICATION NUMBER: 60/215,135
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: 60/225,266
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/249,218
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/249,208
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/249,213
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/249,212
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/249,207
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/249,245
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/249,244
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/249,217
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/249,211
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/249,215
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/249,264
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/249,214
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/249,297

PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/232,400
PRIOR FILING DATE: 2000-09-14
PRIOR APPLICATION NUMBER: 60/231,242
PRIOR FILING DATE: 2000-09-08
PRIOR APPLICATION NUMBER: 60/232,081
PRIOR FILING DATE: 2000-09-08
PRIOR APPLICATION NUMBER: 60/232,080
PRIOR FILING DATE: 2000-09-08
PRIOR APPLICATION NUMBER: 60/231,414
PRIOR FILING DATE: 2000-09-08
PRIOR APPLICATION NUMBER: 60/231,244
PRIOR FILING DATE: 2000-09-08
PRIOR APPLICATION NUMBER: 60/233,064
PRIOR FILING DATE: 2000-09-14
PRIOR APPLICATION NUMBER: 60/233,063
PRIOR FILING DATE: 2000-09-14
PRIOR APPLICATION NUMBER: 60/232,397
PRIOR FILING DATE: 2000-09-14
PRIOR APPLICATION NUMBER: 60/232,399
PRIOR FILING DATE: 2000-09-14
PRIOR APPLICATION NUMBER: 60/241,826
PRIOR FILING DATE: 2000-09-14
PRIOR APPLICATION NUMBER: 60/241,808
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/241,826
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/241,786
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/241,221
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/246,475
PRIOR FILING DATE: 2000-11-08
PRIOR APPLICATION NUMBER: 60/231,243
PRIOR FILING DATE: 2000-09-08

Query Match 39.8%; Score 45; DB 15; Length 58;
Best Local Similarity 56.2%; Pred. No. 11;
Matches 9; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

Qy 1 AIDHSLILEAWSGNDN 16
Db 25 AAGSLYLEGWSGCDH 40

RESULT 10
US-10-011-366-10

Sequence 10, Application US/10011366
Publication No. US20030054493A1

GENERAL INFORMATION:

APPLICANT: Williams, James A.

ATTORNEY/AGENT INFORMATION:

TITLE OF INVENTION: IDENTIFICATION OF NEUTRALIZING EPITOPES

OF TOXIN A AND TOXIN B FOR THE TREATMENT OF C. DIFFICILE

DISEASE

NUMBER OF SEQUENCES: 22

CORRESPONDENCE ADDRESS:

ADDRESSEE: Medlen & Carroll

STREET: 220 Montgomery Street, Suite 2200

CITY: San Francisco

STATE: California

COUNTRY: United States of America

ZIP: 94104

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/10/011,366

FILING DATE: 16-NOV-2003

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/08/957,310
FILING DATE: 23-OCT-1997
APPLICATION NUMBER: US 08/329,154
FILING DATE: 24-OCT-1994
APPLICATION NUMBER: US 08/161,907
FILING DATE: 02-DEC-1993
APPLICATION NUMBER: US 07/985,321
FILING DATE: 04-DEC-1992
APPLICATION NUMBER: US 07/429,791
FILING DATE: 31-OCT-1989
ATTORNEY/AGENT INFORMATION:
NAME: Ingolia, Diane E.
REGISTRATION NUMBER: 40,027
REFERENCE/DOCKET NUMBER: OPND-01121
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 705-8410
TELEFAX: (415) 397-8338
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 2366 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 10:
US-10-011-366-10

Query Match 39.8%; Score 45; DB 15; Length 2366;
Best Local Similarity 44.4%; Pred. No. 66-02;
Matches 8; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

Qy 2 IDHSLILEAWSGNDNDYV 19
Db 1733 INDLSIRYWSNDGNDFI 1750

RESULT 11

US-10-225-810-28

Sequence 28, Application US/10225810

Publication No. US20030157512A1

GENERAL INFORMATION:

APPLICANT: Birmingham, Jr., John R.

TITLE OF INVENTION: Transdorins and Methods of Using Transdorin

FILE REFERENCE: McLaugh-07165

CURRENT APPLICATION NUMBER: US/10/225,810

CURRENT FILING DATE: 2002-08-21

NUMBER OF SEQ ID NOS: 76

SOFTWARE: Patent version 3.2

SEQ ID NO 28

LENGTH: 207

TYPE: PRT

ORGANISM: Homo sapiens

US-10-225-810-28

Query Match 38.9%; Score 44; DB 12; Length 207;
Best Local Similarity 61.5%; Pred. No. 63;
Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 9 EAWSGNDNDYVKQ 21
Db 73 ELWSNDNFTDFVKQ 85

RESULT 12

US-10-225-810-44

Sequence 44, Application US/10225810

Publication No. US20030157512A1

GENERAL INFORMATION:

APPLICANT: Birmingham, Jr., John R.

TITLE OF INVENTION: Transdorins and Methods of Using Transdorin

FILE REFERENCE: McLaugh-07165

CURRENT APPLICATION NUMBER: US/10/225,810

CURRENT FILING DATE: 2002-08-21

NUMBER OF SEQ ID NOS: 76

SOFTWARE: Patentin version 3.2
 SEQ ID NO 44
 LENGTH: 448
 TYPE: PRT
 ORGANISM: Mus musculus
 US-10-225-810-44

Query Match 38.9%; Score 44; DB 12; Length 448;
 Best Local Similarity 61.5%; Pred. No. 1.4e+02;
 Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 9 EAWSGNDNDYVKQ 21
 DB 247 ELWSDNFTDFVKQ 259

RESULT 13
 US-09-810-808-8
 Sequence 8, Application US/09810808
 Patent No. US20020042114A1
 GENERAL INFORMATION:
 APPLICANT: Au-Young, Janice
 Guegler, Karl J.
 Hawkins, Phillip R.

TITLE OF INVENTION: NOVEL HUMAN PROTEIN KINASES
 NUMBER OF SEQUENCES: 9
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Incyte Pharmaceuticals, Inc.
 STREET: 3174 Porter Drive
 CITY: Palo Alto
 STATE: CA
 COUNTRY: U.S.
 ZIP: 94304

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette
 COMPUTER: IBM Compatible
 OPERATING SYSTEM: DOS
 SOFTWARE: FastSeq Version 1.5
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/810,808
 FILING DATE: 15-Mar-2001
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 09/541,228
 FILING DATE: <UNKNOWN>
 ATTORNEY/AGENT INFORMATION:
 NAME: Billings, Lucy J
 REGISTRATION NUMBER: 36,749
 REFERENCE/DOCKET NUMBER: PF-0118 US
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 415-855-0555
 TELEFAX: 415-845-4166

INFORMATION FOR SEQ ID NO: 8:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 487 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: peptide
 IMMEDIATE SOURCE:
 LIBRARY: Genbank
 CLONE: 1117791

US-09-810-808-8
 SEQUENCE DESCRIPTION: SEQ ID NO: 8:

Query Match 38.9%; Score 44; DB 9; Length 487;
 Best Local Similarity 61.5%; Pred. No. 1.6e+02;
 Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 9 EAWSGNDNDYVKQ 21
 DB 248 ELWSDNFTDFVKQ 260

RESULT 14
 US-09-813-408-28
 Sequence 28, Application US/09813408
 Publication No. US20030049619A1
 GENERAL INFORMATION:
 APPLICANT: Delagrave, Simon
 Mairs, Barry

TITLE OF INVENTION: Methods For The Synthesis Of Polynucleotides And Combinatorial Li
 FILE REFERENCE: HER004
 CURRENT APPLICATION NUMBER: US/09/813,408
 CURRENT FILING DATE: 2001-03-21
 NUMBER OF SEQ ID NOS: 85
 SOFTWARE: Patentin version 3.0
 SEQ ID NO 28
 LENGTH: 554
 TYPE: PRT
 ORGANISM: Bacillus halodurans
 US-09-813-408-28

Query Match 38.9%; Score 44; DB 11; Length 554;
 Best Local Similarity 62.5%; Pred. No. 1.8e+02;
 Matches 10; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 AIDHSLIEAWSGNDN 16
 DB 450 AYDHDVILIAASGNDN 465

RESULT 15
 US-10-291-253A-11
 Sequence 11, Application US/10291253A
 Publication No. US20030150017A1
 GENERAL INFORMATION:
 APPLICANT: Botella, Jose
 Graham, Michael

APPLICANT: Fairbairn, David
 TITLE OF INVENTION: A Method For Facilitating Pathogen Resistance
 FILE REFERENCE: nematode
 CURRENT APPLICATION NUMBER: US/10/291,253A
 CURRENT FILING DATE: 2003-03-31
 PRIOR APPLICATION NUMBER: PR8706
 PRIOR FILING DATE: 2001-11-07
 PRIOR APPLICATION NUMBER: PR8802
 PRIOR FILING DATE: 2001-11-12
 PRIOR APPLICATION NUMBER: US60/341404
 PRIOR FILING DATE: 2001-12-14
 NUMBER OF SEQ ID NOS: 16
 SOFTWARE: Patentin version 3.1
 SEQ ID NO 11
 LENGTH: 272
 TYPE: PRT
 ORGANISM: Meloidogyne javanica
 US-10-291-253A-11

Query Match 38.1%; Score 43; DB 12; Length 272;
 Best Local Similarity 35.7%; Pred. No. 1.2e+02;
 Matches 10; Conservative 3; Mismatches 7; Indels 8; Gaps 1;

QY 2 IDHSLIEAWSGNDNDYVKQ 21
 DB 177 VDDSLIVKRWVSKSKFSNEGNDREKQ 204

Search completed: November 13, 2003, 10:29:03
 Job time : 27.4692 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 13, 2003, 09:31:40 ; Search time 38.5166 Seconds
(without alignments)
86.541 Million cell updates/sec

Title:	US-09-290-049A-14
Perfect score:	119
Sequence:	1 ANNHVSIVEAWSNDNDTPYLHD 21

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

```
Searched:      1107863 segs, 158726573 residues
Total number of hits satisfying chosen parameters: 1107863
```

```
Minimum DB seq length: 0
Maximum DB seq length: 20000000000
```

Post-processing:	Minimum Match	0%
	Maximum Match	100%
	Listing first	45 summaries

A_Geneseq_19Jun03: *

1:	/SIDS1/gcgdata/geneseq/geneseq-emb1/AA1980.DAT: *
2:	/SIDS1/gcgdata/geneseq/geneseq-emb1/AA1981.DAT: *
3:	/SIDS1/gcgdata/geneseq/geneseq-emb1/AA1982.DAT: *
4:	/SIDS1/gcgdata/geneseq/geneseq-emb1/AA1983.DAT: *
5:	/SIDS1/gcgdata/geneseq/geneseq-emb1/AA1984.DAT: *
6:	/SIDS1/gcgdata/geneseq/geneseq-emb1/AA1985.DAT: *
7:	/SIDS1/gcgdata/geneseq/geneseq-emb1/AA1986.DAT: *
8:	/SIDS1/gcgdata/geneseq/geneseq-emb1/AA1987.DAT: *
9:	/SIDS1/gcgdata/geneseq/geneseq-emb1/AA1988.DAT: *
10:	/SIDS1/gcgdata/geneseq/geneseq-emb1/AA1989.DAT: *
11:	/SIDS1/gcgdata/geneseq/geneseq-emb1/AA1990.DAT: *
12:	/SIDS1/gcgdata/geneseq/geneseq-emb1/AA1991.DAT: *
13:	/SIDS1/gcgdata/geneseq/geneseq-emb1/AA1992.DAT: *
14:	/SIDS1/gcgdata/geneseq/geneseq-emb1/AA1993.DAT: *
15:	/SIDS1/gcgdata/geneseq/geneseq-emb1/AA1994.DAT: *
16:	/SIDS1/gcgdata/geneseq/geneseq-emb1/AA1995.DAT: *
17:	/SIDS1/gcgdata/geneseq/geneseq-emb1/AA1996.DAT: *
18:	/SIDS1/gcgdata/geneseq/geneseq-emb1/AA1997.DAT: *
19:	/SIDS1/gcgdata/geneseq/geneseq-emb1/AA1998.DAT: *
20:	/SIDS1/gcgdata/geneseq/geneseq-emb1/AA1999.DAT: *
21:	/SIDS1/gcgdata/geneseq/geneseq-emb1/AA2000.DAT: *
22:	/SIDS1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT: *
23:	/SIDS1/gcgdata/geneseq/geneseq-emb1/AA2002.DAT: *
24:	/SIDS1/gcgdata/geneseq/geneseq-emb1/AA2003.DAT: *

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	119	100.0	1552	14	AAU32925	Glucosyltransferase
2	108	90.8	1017	23	AAU79285	Streptococcus mutans
3	108	90.8	1475	23	AAU98027	S. mutans glucosyl
4	108	90.8	1475	23	AAU98030	S. mutans glucosyl
5	108	90.8	1475	23	AAU98031	S. mutans glucosyl
6	108	90.8	1475	23	AAU98032	S. mutans glucosyl
7	108	90.8	1475	23	AAU98033	S. mutans glucosyl
8	108	90.8	1475	23	AAU98034	S. mutans glucosyl
9	108	90.8	1475	23	AAU98035	S. mutans glucosyl

10	108	90.8	1475	23	AAU98036
11	108	90.8	1475	23	AAU98037
12	108	90.8	1475	23	AAU98038
13	108	90.8	1475	23	AAU98039
14	108	90.8	1475	23	AAU98040
15	108	90.8	1475	23	AAU972844
16	99	83.2	1375	23	AAU980288
17	99	83.2	1375	23	AAU972888
18	74	62.2	1430	23	AAU980299
19	74	62.2	1430	23	AAU98041
20	74	62.2	1430	23	AAU98042
21	74	62.2	1430	23	AAU98043
22	74	62.2	1430	23	AAU98044
23	74	62.2	1430	23	AAU98045
24	72	60.5	285	23	ABB98574
25	66	55.6	152	23	ABB98642
26	65	54.5	157	23	AAU80055
27	60	50.4	172	23	AAU80641
28	57	47.9	1577	21	AAU81047
29	55	46.2	2057	21	AAU80657
30	53	44.5	195	21	ABB98643
31	53	44.5	195	21	AAU90941
32	52	43.7	895	23	ABB98573
33	51	42.9	1195	20	AAU75420
34	50	42.0	1222	22	ABB98363
35	48.5	40.8	441	22	AAU35544
36	48	40.3	1781	23	AAU74519
37	47	39.5	112	23	ABB98678
38	47	39.5	112	23	ABB98644
39	47	39.5	12	23	ABB98646
40	47	39.5	12	23	ABB98648
41	47	39.5	759	23	ABB04873
42	46.5	39.1	1604	23	AAU87244
43	46.5	39.1	1604	23	AAU82715
44	46	38.7	486	22	AAU36552
45	46	38.7	880	22	ABB61809

ALIGNMENTS

RESULT 1	
AA32925	
ID	AA32925 strand; Protein; 1592 AA.
XX	
AC	AA32925;
XX	
DT	28-JUN-1993 (first entry)
XX	
DE	Glucosyltransferase I.
XX	
KW	GT-1; Streptococcus; dental; caries.
XX	
OS	Streptococcus sobrinus.
XX	
PN	JP05023188-A.
XX	
PD	02-FEB-1993.
XX	
PF	25-JUL-1991; 91UP-0186592.
XX	
PR	25-JUL-1991; 91UP-0186592.
XX	
PA	(FUKU/) FUKUI I.
PA	(KATO/) KATO K.
XX	
DR	WPI; 1993-079449/10.
XX	
PT	N-PSDB; AAQ37760.
XX	
PT	DNA sequence glucosyl:transferase-I - comprises Streptococcus
XX	sobrinus DNA sequence with at least one nucleotide added or
XX	deleted
XX	

PS Claim 13; Page 15; 29pp; Japanese.
XX
CC The DNA sequence from Streptococcus sobrinus strain 6715 encodes
CC glucosyltransferase-I (and mutants). The DNA was obt. by treating
CC S. sobrinus 6715 with mutanolysin, extracting the chromosomal DNA,
CC partially digesting with Sau3AI and fractionating on agarose gel.
CC The 3-5 kbp fragment was ligated into pUC18 and E. coli DH109
CC transformed with it. A GT-1 expressing clone was isolated and
CC sequenced. The clone may be used in the development of a drug for
CC dental caries.
XX
SQ Sequence 1592 AA;
Query Match 100.0%; Score 119; DB 14; Length 1592;
Best Local Similarity 100.0%; Pred. No. 6.4e-09;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ANNHVSIVEAWSDNDTPYLHD 21
Db 477 ANNHVSIVEAWSDNDTPYLHD 497
RESULT 2
AAU9285
ID AAU9285 standard; Protein; 1017 AA.
XX
AC AAU9285;
XX
DT 13-AUG-2002 (first entry)
XX
DE Streptococcus mutans monoclonal antibody-related protein #2.
XX
KW Antibody; dental caries; water insoluble glucan synthetase;
KW anti-carries; glucosyl transferase-B; immunotherapy.
XX
OS Streptococcus mutans.
XX
PN JP2002114709-A.
XX
PD 16-APR-2002.
XX
PF 04-OCT-2000; 2000JP-0304889.
XX
PR 04-OCT-2000; 2000JP-0304889.
XX
PS (UYN1-) UNIT NIPON.
XX
DR WPI; 2002-448101/48.
XX
PT Anti-carries agent composed of a monoclonal antibody against an
PT inhibitory enzyme against water insoluble glucan synthetase of glucosyl
PT transferase-B (GTF-B) of Streptococcus mutans -
XX
PS Claim 4; Page 17-19; 28pp; Japanese.
XX
CC The invention relates to a monoclonal antibody against dental caries and
CC an anti-carries agent composed of a monoclonal antibody produced by
CC Streptococcus mutans, particularly mouse-hybridoma MHP126 (FERM P-17566)
CC or mouse-hybridoma MHP136 (FERM P-17567), against an enzyme having
CC inhibitive activity against water insoluble glucan synthetase of glucosyl
CC transferase-B. The monoclonal antibody specifically inhibits water
CC insoluble glucan synthetase of Streptococcus mutans produced glucosyl
CC transferase-B and is used in the immunotherapy of dental caries. This
CC sequence represents a Streptococcus mutans monoclonal antibody-related
CC protein.
XX
SQ Sequence 1017 AA;
Query Match 90.8%; Score 108; DB 23; Length 1017;
Best Local Similarity 85.7%; Pred. No. 1.7e-07;
Matches 18; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
QY 1 ANNHVSIVEAWSDNDTPYLHD 21

Db 447 ANNHVSIVEAWSDNDTPYLHD 467
RESULT 3
AAU98027
ID AAU98027 standard; Protein; 1475 AA.
XX
AC AAU98027;
XX
DT 27-AUG-2002 (first entry)
XX
DE S. mutans glucosyltransferase GTFB.
XX
KW Glucosyltransferase; GTFB; transgenic plant; paper sizing;
KW coating composition; glucan; starch; latex; thermoplastic molecule;
KW amyloplast; vacuole; paper manufacture.
XX
OS Streptococcus mutans.
XX
PN US2002031826-A1.
XX
PD 14-MAR-2002.
XX
PF 19-DEC-2000; 2000US-0740274.
XX
PR 11-DEC-1998; 98US-0210361.
XX
PR 07-JUN-1995; 95US-0478704.
XX
PR 07-JUN-1995; 95US-0482711.
XX
PR 07-JUN-1995; 95US-0485243.
XX
PR 16-JAN-1998; 98US-0007999.
XX
PR 16-JAN-1998; 98US-0008172.
XX
PR 20-JAN-1998; 98US-0009620.
XX
PA (NICH/) NICHOLS S E.
XX
PI Nichols SE;
XX
DR WPI; 2002-414332/44.
XX
DR N-PSDB; ABR52938.
XX
PT Glucosyltransferase B or D protein useful for producing a glucan useful
PT as substitutes for and additions to modified starch and latexes in
PT paper manufacture, comprises mutations in specific positions -
XX
PS Disclosure; Page 21-25; 44pp; English.
XX
CC The invention an isolated protein comprising a glucosyltransferase
CC (GTF) B polypeptide having changes at position from 1448V, D457N,
CC D567T, K1014T, D457N/D567T, D457N/D571K, D567T/D571K,
CC D567T/D571K/K1014T, I448V/D457N/D567T/D571K/K799Q/K1014T,
CC Y163N/Y170K/Y171A, and K779Q or a GTF D polypeptide having
CC changes at positions from T589D, T589S, N471D, N471D/T589D, and
CC N471D/T589S. Also included are a glucan produced by the GTF mutant,
CC an isolated polynucleotide which encodes P1 or P2, or its complementary
CC polynucleotide, a ribonucleic acid sequence encoding the GTF mutant,
CC an expression cassette comprising the polynucleotide operably linked to a
CC promoter, a vector comprising the expression cassette, host cell
CC introduced with the vector, a transgenic plant comprising the
CC vector, a seed or tuber from the transgenic plant, a paper sizing and/or
CC coating composition comprising a glucan produced in a plant transformed
CC with a gene encoding the mutant GTF, wild type or, starch, a latex,
CC thermoplastic molecule or their combinations or glucan and starch where
CC the glucan is produced in the amyloplast and/or vacuole or a maize line
CC deficient in starch biosynthesis, transformed with a gene encoding a
CC glucosyltransferase B or D enzyme, wild-type or mutant and a paper
CC comprising the glucan (paper sizing/coating agent). The vector is useful
CC for producing a glucan in a plant. The method comprises transforming a
CC plant cell with the vector, growing the plant cell under plant growing
CC conditions to produce a regenerated plant and inducing expression of the
CC polynucleotide for a time sufficient to produce the glucan in the
CC regenerated plant, where the vector contains a transit sequence from
CC ribulose biphosphate carboxylase small subunit, waxy, brittle-1 and

	CC	chlorophyll AB binding protein to produce a transgenic plant, and glucan
	CC	is produced in the amyloplast of potato or the vacuole of sugar beet.
	CC	Glucans are useful as substitutes for and additions to modified starch.
	CC	and latexes in paper manufacture. Unlike prior art techniques, which
	CC	require input materials that produce chemical effluents, paper
	CC	manufacture utilizing the glucan produced by GTFB, which utilizes
	CC	biologically produced input materials, is more cost-effective and
	CC	environmentally friendly. Moreover, glucans also exhibit thermoplastic
	CC	properties and impart gloss to the paper during coating step.
	CC	The present sequence represents GTFB.
	XX	
	SQ	Sequence 1475 AA;
	CY	
	D8	
	DB	
	Query Match	90.8%; Score 108; DB 23; Length 1475;
	Best Local Similarity	85.7%; Pred. No. 2.7e-07;
	Matches 18; Conservative 3; Mismatches 0; Indels 0; Gaps 0;	
	1 ANNHVSTVEASDNDTPYLHD 21	
	::: :	
	481 ANDHLSIDEAWSDNDTPYLHD 501	
	RESULT 4	
	AU98030	
	ID AU98030 standard; Protein, 1475 AA.	
	XX AAU98030;	
	AC	
	XX 27-AUG-2002 (first entry)	
	DT	
	DE S. mutans glucosyltransferase GTFB mutant I448V.	
	XX Glucosyltransferase; GTFB; transgenic plant; paper sizing;	
	KW coating composition; glucan; starch; latex; thermoplastic molecule;	
	KM amyloplastic; vacuole; paper manufacture; mutant; mutein.	
	XX Streptococcus mutans.	
	OS Synthetic.	
	XX Key Location/Qualifiers	
	FH Misc-difference 448 /note= "Wild-type Ile substituted by Val"	
	FT US2002031826-A1.	
	PN	
	XX 14-MAR-2002.	
	PD	
	PF 19-DEC-2000; 2000US-0740274.	
	XX 11-DEC-1998; 98US-0210361.	
	PR 07-JUN-1995; 95US-0478704.	
	PR 07-JUN-1995; 95US-0482711.	
	PR 07-JUN-1995; 95US-0485243.	
	PR 16-JAN-1998; 98US-0007399.	
	PR 16-JAN-1998; 98US-0008172.	
	PR 20-JAN-1998; 98US-0009620.	
	PA (NICH/) NICHOLS S E.	
	PI Nichols SE;	
	DR WPT, 2002-414332/44.	
	XX Glucosyltransferase B or D protein useful for producing a glucan useful	
	PT as substitutes for and additions to modified starch and latexes in	
	PT paper manufacture, comprises mutations in specific positions -	
	XX Claim 36; Page -, 44pp; English.	
	XX The invention of an isolated protein comprising a glucosyltransferase	
	CC (GTF) B polypeptide having changes at position from I448V, D457N,	
	CC D567T, K1014T, D457N/D567T, D457N/D571K, D567T/D571K,	
	CC D567T/D571K/K1014T, I448V/D457N/D567T/D571K/K739Q/K1014T,	

[illegible]

XX 11-DEC-1998; 98US-0210361.
 PR 07-JUN-1995; 95US-0478704.
 PR 07-JUN-1995; 95US-0482711.
 PR 07-JUN-1995; 95US-0485243.
 PR 16-JAN-1998; 98US-0007999.
 PR 16-JAN-1998; 98US-0008172.
 PR 20-JAN-1998; 98US-0009620.
 XX
 PA (NICH/) NICHOLS S E.
 PI Nichols SE;
 DR WPI; 2002-414332/44.
 XX
 PT Glucosyltransferase B or D protein useful for producing a glucan useful
 as substitutes for and additions to modified starch and latexes in
 PT paper manufacture, comprises mutations in specific positions -
 XX
 PS Claim 36; Page -; 44pp; English.

XX The invention an isolated protein comprising a glucosyltransferase
 CC (GTF) B polypeptide having changes at position from 1448V, D457N,
 CC D567T, K1014T, D457N/D567T, D457N/D571K, D567T/D571K,
 CC D567T/D571K/K1014T, 1448V/D457N/D567T/D571K/K779Q/K1014T,
 CC Y169A/Y170A/Y171A, and K779Q or a GTF D polypeptide having
 CC changes at positions from T589D, T589E, N471D, N471D/T589D, and
 CC N471D/T589E. Also included are a glucan produced by the GTF mutant,
 CC an isolated polynucleotide which encodes P1 or P2, or its complementary
 CC polynucleotide, a ribonucleic acid sequence encoding the GTF mutant,
 CC an expression cassette comprising the polynucleotide operably linked to a
 CC promoter, a vector comprising the expression cassette, host cell
 CC introduced with the vector, a transgenic plant comprising the
 CC vector, a seed or tuber from the transgenic plant, a paper sizing and/or
 CC coating composition comprising a glucan produced in a plant transformed
 CC with a gene encoding the mutant GTF, wild type or, starch, a latex,
 CC thermoplastic molecule or their combinations or glucan and starch where
 CC the glucan is produced in the amyloplast and/or vacuole or a maize line
 CC deficient in starch biosynthesis, transformed with a gene encoding a
 CC glucosyltransferase B or D enzyme, wild-type or mutant and a paper
 CC comprising the glucan (paper sizing/coating agent). The vector is useful
 CC for producing a glucan in a plant. The method comprises transforming a
 CC plant cell with the vector, growing the plant cell under plant growing
 CC conditions to produce a regenerated plant and inducing expression of the
 CC polynucleotide for a time sufficient to produce the glucan in the
 CC regenerated plant, where the vector contains a transit sequence from
 CC ribulose biphosphate carboxylase small subunit, waxy, brittle-1 and
 CC chlorophyll AB binding protein to produce a transgenic plant, and glucan
 CC is produced in the amyloplast of potato or the vacuole of sugar beet.
 CC Glucans are useful as substitutes for and additions to modified starch
 CC and latexes in paper manufacture. Unlike prior art techniques, which
 CC require input materials that produce chemical effluents, paper
 CC manufacture utilizing the glucan produced by GTF, which utilizes
 CC biologically produced input materials, is more cost-effective and
 CC environmentally friendly. Moreover, glucans also exhibit thermoplastic
 CC properties and impart gloss to the paper during coating step.
 CC The present sequence represents a GTFB mutant of the invention.
 CC Note: The present sequence is not shown in the specification but
 CC was created by the indexer using the GTFB sequence appearing as AAU98027
 CC and the information in claim 36.

XX Sequence 1475 AA;
 SQ
 Query Match 90.8%; Score 108; DB 23; Length 1475;
 Best Local Similarity 85.7%; Pred. No. 2.7e-07;
 Matches 18; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 ANNVSVLEAASDNDTPTLHD 21
 Db 481 ANNDLSILAEASDNDTPTLHD 501

RESULT 6

AAU98032
 ID AAU98032 standard; Protein; 1475 AA.
 XX
 AC AAU98032;
 XX
 DT 27-AUG-2002 (first entry)
 XX
 DE 5. mutans glucosyltransferase GTFB mutant D567T.
 XX
 KW Glucosyltransferase; GTFB; transgenic plant; paper sizing;
 KW coating composition; glucan; starch; latex; thermoplastic molecule;
 KW amyloplast; vacuole; paper manufacture; mutant; mutin.
 XX
 OS Streptococcus mutans.
 OS Synthetic.
 XX
 FH Key Location/Qualifiers
 FT Misc-difference 567
 FT /note= "Wild-type Asp substituted by Thr"

XX US2002031826-A1.
 XX
 PD 14-MAR-2002.
 XX
 PF 19-DEC-2000; 2000US-0740274.
 XX
 XX 11-DEC-1998; 98US-0210361.
 PR 07-JUN-1995; 95US-0478704.
 PR 07-JUN-1995; 95US-0482711.
 PR 07-JUN-1995; 95US-0485243.
 PR 16-JAN-1998; 98US-0007999.
 PR 16-JAN-1998; 98US-0008172.
 PR 20-JAN-1998; 98US-0009620.
 XX
 PA (NICH/) NICHOLS S E.
 XX
 PI Nichols SE;
 DR WPI; 2002-414332/44.
 XX
 PT Glucosyltransferase B or D protein useful for producing a glucan useful
 as substitutes for and additions to modified starch and latexes in
 PT paper manufacture, comprises mutations in specific positions -
 XX
 PS Claim 36; Page -; 44pp; English.

XX The invention an isolated protein comprising a glucosyltransferase
 CC (GTF) B polypeptide having changes at position from 1448V, D457N,
 CC D567T, K1014T, D457N/D567T, D457N/D571K, D567T/D571K,
 CC D567T/D571K/K1014T, 1448V/D457N/D567T/D571K/K779Q/K1014T,
 CC Y169A/Y170A/Y171A, and K779Q or a GTF D polypeptide having
 CC changes at positions from T589D, T589E, N471D, N471D/T589D, and
 CC N471D/T589E. Also included are a glucan produced by the GTF mutant,
 CC an isolated polynucleotide which encodes P1 or P2, or its complementary
 CC polynucleotide, a ribonucleic acid sequence encoding the GTF mutant,
 CC an expression cassette comprising the polynucleotide operably linked to a
 CC promoter, a vector comprising the expression cassette, host cell
 CC introduced with the vector, a transgenic plant comprising the
 CC vector, a seed or tuber from the transgenic plant, a paper sizing and/or
 CC coating composition comprising a glucan produced in a plant transformed
 CC with a gene encoding the mutant GTF, wild type or, starch, a latex,
 CC thermoplastic molecule or their combinations or glucan and starch where
 CC the glucan is produced in the amyloplast and/or vacuole or a maize line
 CC deficient in starch biosynthesis, transformed with a gene encoding a
 CC glucosyltransferase B or D enzyme, wild-type or mutant and a paper
 CC comprising the glucan (paper sizing/coating agent). The vector is useful
 CC for producing a glucan in a plant. The method comprises transforming a
 CC plant cell with the vector, growing the plant cell under plant growing
 CC conditions to produce a regenerated plant and inducing expression of the
 CC polynucleotide for a time sufficient to produce the glucan in the
 CC regenerated plant, where the vector contains a transit sequence from
 CC ribulose biphosphate carboxylase small subunit, waxy, brittle-1 and
 CC chlorophyll AB binding protein to produce a transgenic plant, and glucan

XX 14-MAR-2002.
 PD 19-DEC-2000; 2000US-0740274.
 PF 11-DEC-1998; 98US-0210361.
 XX 07-JUN-1995; 95US-0478704.
 PR 07-JUN-1995; 95US-0482711.
 PR 07-JUN-1995; 95US-0485243.
 PR 16-JAN-1998; 98US-0007999.
 PR 16-JAN-1998; 98US-0008172.
 XX 20-JAN-1998; 98US-0009620.
 XX (NICH/) NICHOLS S E.
 PI Nichols SE;
 DR WPI; 2002-414332/44.
 XX
 PT Glucosyltransferase B or D protein useful for producing a glucan useful
 PT as substitutes for and additions to modified starch and latexes in
 PT paper manufacture, comprises mutations in specific positions -
 PS Claim 36; Page -; 44pp; English.
 XX
 CC The invention an isolated protein comprising a glucosyltransferase
 CC (GTF) B polypeptide having changes at position from 1448V, D457N,
 CC D567T, K1014T, D457N/D567T, D457N/D571K, D567T/D571K,
 CC D567T/D571K/K1014T, I448V/D457N/D567T/D571K/K779Q/K1014T,
 CC Y169A/Y170A/Y171A, and K779Q or a GTF D polypeptide having
 CC changes at positions from T589D, T589E, N471D, N471D/T589D, and
 CC N471D/T589E. Also included are a glucan produced by the GTF mutant,
 CC an isolated polynucleotide which encodes P1 or P2, or its complementary
 CC polynucleotide, a ribonucleic acid sequence encoding the GTF mutant,
 CC an expression cassette comprising the polynucleotide operably linked to a
 CC promoter, a vector comprising the expression cassette, host cell
 CC introduced with the vector, a transgenic plant comprising the
 CC vector, a seed or tuber from the transgenic plant, a paper sizing and/or
 CC coating composition comprising a glucan produced in a plant transformed
 CC with a gene encoding the mutant GTF, wild type or, starch, a latex,
 CC thermoplastic molecule or their combinations or glucan and starch where
 CC the glucan is produced in the amyloplast and/or vacuole of a maize line
 CC deficient in starch biosynthesis, transformed with a gene encoding a
 CC glucosyltransferase B or D enzyme, wild-type or mutant and a paper
 CC comprising the glucan (paper sizing/coating agent). The vector is useful
 CC for producing a glucan in a plant. The method comprises transforming a
 CC plant cell with the vector, growing the plant cell under plant growing
 CC conditions to produce a regenerated plant and inducing expression of the
 CC polynucleotide for a time sufficient to produce the glucan in the
 CC regenerated plant, where the vector contains a transit sequence from
 CC ribulose biphosphate carboxylase small subunit, waxy, brittle-1 and
 CC chlorophyll AB binding protein to produce a transgenic plant, and glucan
 CC is produced in the amyloplast of potato or the vacuole of sugar beet.
 CC Glucans are useful as substitutes for and additions to modified starch
 CC and latexes in paper manufacture. Unlike prior art techniques, which
 CC require input materials that produce chemical effluents, paper
 CC manufacture utilizing the glucan produced by GTF, which utilizes
 CC biologically produced input materials, is more cost-effective and
 CC environmentally friendly. Moreover, glucans also exhibit thermoplastic
 CC properties and impart gloss to the paper during coating step.
 CC The present sequence represents a GTFB mutant of the invention.
 CC Note: The present sequence is not shown in the specification but
 CC was created by the indexer using the GTFB sequence appearing as AAU98027
 CC and the information in claim 36.
 XX
 XX Sequence 1475 AA;
 SQ
 Query Match 90.8%; Score 108; DB 23; Length 1475;
 Best Local Similarity 85.7%; Pred. No. 2,7e-07;
 Matches 18; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
 QY 1 ANNHVSGIVEAKMSDNDPTPLMD 21
 ||:|||||

DB 481 ANHDLSTLEAKMSDNDPTPLMD 501
 RESULT 9
 ID AAU98035 standard; Protein; 1475 AA.
 XX AAU98035;
 AC
 XX 27-AUG-2002 (first entry)
 DT
 XX
 DE S. mutans glucosyltransferase GTFB mutant D457N/D571K.
 XX
 KW Glucosyltransferase; GTFB; transgenic plant; paper sizing;
 KW coating composition; glucan; starch; latex; thermoplastic molecule;
 KW amyloplast; vacuole; paper manufacture; mutant; mutain.
 XX
 OS Streptococcus mutans.
 OS Synthetic.
 XX
 FH Key Location/Qualifiers
 FT Misc-difference 457
 FT /note= "Wild-type Asp substituted by Asn"
 FT Misc-difference 571
 FT /note= "Wild-type Asp substituted by Lys"
 FT
 XX US2002031826-A1.
 PN
 XX 14-MAR-2002.
 PD
 XX 19-DEC-2000; 2000US-0740274.
 PF
 XX 11-DEC-1998; 98US-0210361.
 PR 07-JUN-1995; 95US-0478704.
 PR 07-JUN-1995; 95US-0482711.
 PR 07-JUN-1995; 95US-0485243.
 PR 16-JAN-1998; 98US-0007999.
 PR 16-JAN-1998; 98US-0008172.
 PR 20-JAN-1998; 98US-0009620.
 XX
 PA (NICH/) NICHOLS S E.
 PI Nichols SE;
 DR WPI; 2002-414332/44.
 XX
 PT Glucosyltransferase B or D protein useful for producing a glucan useful
 PT as substitutes for and additions to modified starch and latexes in
 PT paper manufacture, comprises mutations in specific positions -
 PS Claim 36; Page -; 44pp; English.
 XX
 CC The invention an isolated protein comprising a glucosyltransferase
 CC (GTF) B polypeptide having changes at position from 1448V, D457N,
 CC D567T, K1014T, D457N/D567T, D457N/D571K, D567T/D571K,
 CC D567T/D571K/K1014T, I448V/D457N/D567T/D571K/K779Q/K1014T,
 CC Y169A/Y170A/Y171A, and K779Q or a GTF D polypeptide having
 CC changes at positions from T589D, T589E, N471D, N471D/T589D, and
 CC N471D/T589E. Also included are a glucan produced by the GTF mutant,
 CC an isolated polynucleotide which encodes P1 or P2, or its complementary
 CC polynucleotide, a ribonucleic acid sequence encoding the GTF mutant,
 CC an expression cassette comprising the polynucleotide operably linked to a
 CC promoter, a vector comprising the expression cassette, host cell
 CC introduced with the vector, a transgenic plant comprising the
 CC vector, a seed or tuber from the transgenic plant, a paper sizing and/or
 CC coating composition comprising a glucan produced in a plant transformed
 CC with a gene encoding the mutant GTF, wild type or, starch, a latex,
 CC thermoplastic molecule or their combinations or glucan and starch where
 CC the glucan is produced in the amyloplast and/or vacuole of a maize line
 CC deficient in starch biosynthesis, transformed with a gene encoding a
 CC glucosyltransferase B or D enzyme, wild-type or mutant and a paper
 CC comprising the glucan (paper sizing/coating agent). The vector is useful
 CC for producing a glucan in a plant. The method comprises transforming a


```

CC plant cell with the vector, growing the plant cell under plant growing
CC conditions to produce a regenerated plant and inducing expression of the
CC polynucleotide for a time sufficient to produce the glucan in the
CC regenerated plant, where the vector contains a transit sequence from
CC ribulose biphosphate carboxylase small subunit, waxy, brittle-1 and
CC chlorophyll a/b binding protein to produce a transgenic plant, and glucan
CC is produced in the amyloplast of potato or the vacuole of sugar beet.
CC Glucans are useful as substitutes for and additions to modified starch
CC and latexes in paper manufacture. Unlike prior art techniques, which
CC require input materials that produce chemical effluents, paper
CC manufacture utilizing the glucan produced by GFP, which utilizes
CC biologically produced input materials, is more cost-effective and
CC environmentally friendly. Moreover, glucans also exhibit thermoplastic
CC properties and impart gloss to the paper during coating step.
CC The present sequence represents a GFPB mutant of the invention.
CC Note: The present sequence is not shown in the specification but
CC was created by the indexer using the GFPB sequence appearing as AAU98027
CC and the information in claim 36.
CC
CC
CC
SQ Sequence 1475 AA;

Query Match 90.8%; Score 108; DB 23; Length 1475;
Best Local Similarity 85.7%; Pred. No. 2.7e-07;
Matches 18; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Dy 1 ANNHVSIVEAMSDNDTPYLAD 21
   ||:|:|:|:|:|:|:|:|:|
Db 481 ANDHSLIEAMSDNDTPYLHD 501

RESULT 10
AAU98036
ID AAU98036 standard; Protein; 1475 AA.
XX
XX AAU98036;
XX
XX 27-AUG-2002 (first entry)
XX
XX
XX 5. mutans glucosyltransferase GFPB mutant D5677/D571K.
DE
XX
XX Glucosyltransferase; GFPB; transgenic plant; paper sizing;
KW coating composition; glucan; starch; latex; thermoplastic molecule;
KW amyloplast; vacuole; paper manufacture; mutant; mutain.
XX
XX Streptococcus mutans.
OS Synthetic.
XX
XX Key Location/Qualifiers
FH Misc-difference 567
FT /note= "Wild-type Asp substituted by Thr"
FT Misc-difference 571
FT /note= "Wild-type Asp substituted by Lys"
XX
XX US2002031826-A1.
PN
XX
XX 14-MAR-2002.
PD
XX
XX 19-DEC-2000; 2000US-0740274.
XX
XX 11-DEC-1998; 98US-0210361.
XX 07-JUN-1995; 95US-0478704.
XX 07-JUN-1995; 95US-0482711.
XX 07-JUN-1995; 95US-0485243.
XX 16-JAN-1998; 98US-0007999.
XX 16-JAN-1998; 98US-0008172.
XX 20-JAN-1998; 98US-0009620.
XX
XX (NICH/) NICHOLS S E.
XX
XX Nichols SE;
XX
XX WPI; 2002-414332/44.
XX

```

Claim 36; Page -; 44pp; English.

The invention an isolated protein comprising a glucosyltransferase (GTF) B polypeptide having changes at position from 1448V, D457N, D567T, K1014T, D457N/D567T, D457N/D571K, D567T/D571K, D567T/D571K/K1014T, 1448V/D457N/D567T/D571K/K779Q/K1014T, Y169N/Y170A/Y171A, and K779Q or a GTF D polypeptide having changes at positions from T589D, T589E, N471D, N471D/T589D, and N471D/T589E. Also included are a glucan produced by the GTF mutant, an isolated polynucleotide which encodes P1 or P2, or its complementary polynucleotide, a ribonucleic acid sequence encoding the GTF mutant, an expression cassette comprising the polynucleotide operably linked to a promoter, a vector comprising the expression cassette, host cell introduced with the vector, a transgenic plant comprising the vector, a seed or tuber from the transgenic plant, a paper sizing and/or coating composition comprising a glucan produced in a plant transformed with a gene encoding the mutant GTF, wild type or starch, a latex, thermoplastic molecule or their combinations or glucan and starch where the glucan is produced in the amyloplast and/or vacuole or a maize line deficient in starch biosynthesis, transformed with a gene encoding a glucosyltransferase B or D enzyme, wild-type or mutant and a paper comprising the glucan (paper sizing/coating agent). The vector is useful for producing a glucan in a plant. The method comprises transforming a plant cell with the vector, growing the plant cell under plant growing conditions to produce a regenerated plant and inducing expression of the polynucleotide for a time sufficient to produce the glucan in the regenerated plant, where the vector contains a transit sequence from ribulose biphosphate carboxylase small subunit, waxy, brittle-1 and chlorophyll a binding protein to produce a transgenic plant, and glucan is produced in the amyloplast of potato or the vacuole of sugar beet. Glucans are useful as substitutes for and additions to modified starch and latexes in paper manufacture. Unlike prior art techniques, which require input materials that produce chemical effluents, paper manufacture utilizing the glucan produced by GTF, which utilizes biologically produced input materials, is more cost-effective and environmentally friendly. Moreover, glucans also exhibit thermoplastic properties and impart gloss to the paper during coating step. The present sequence represents a GTFB mutant of the invention. Note: The present sequence is not shown in the specification but was created by the indexer using the GTFB sequence appearing as AA098027 and the information in Claim 36.

Sequence 1475 AA;

Query Match 50.8%; Score 108; DB 23; Length 1475;
Best Local Similarity 85.7%; Pred. No. 2.7e-07.
Matches 18; Conservative 3; Mismatches 0; Indels 0; Gaps 0.

QY 1 ANNHVSIEAWSDNDTPYIHD 21
DB 481 ANHDISTIEAWSDNDTPYIHD 501

RESULT 11
AA098037
ID AA098037 standard; Protein; 1475 AA.
AC AA098037;
XX
XX
DT 27-AUG-2002 (first entry)

5. mutans glucosyltransferase GTFB mutant D567T/D571K/K1014T.
XX
XX
XX Glucosyltransferase; GTFB; transgenic plant; paper sizing;
KM coating composition; glucan; starch; latex; thermoplastic molecule;
KM amyloplast; vacuole; paper manufacture; mutant; mutain.
XX
XX Streptococcus mutans.
OS Synthetic.

XX Key Location/Qualifiers
 FH Misc-difference 567 /note= "Wild-type Asp substituted by Thr"
 FT Misc-difference 571 /note= "Wild-type Asp substituted by Lys"
 FT Misc-difference 1014 /note= "Wild-type Lys substituted by Thr"
 FT US2002031826-A1.
 PN 14-MAR-2002.
 PD 19-DEC-2000; 2000US-0740274.
 XX 11-DEC-1998; 98US-0210361.
 XX 07-JUN-1995; 95US-0478704.
 PR 07-JUN-1995; 95US-0482711.
 PR 07-JUN-1995; 95US-0485243.
 PR 16-JAN-1998; 98US-0007999.
 PR 16-JAN-1998; 98US-0008172.
 PR 20-JAN-1998; 98US-0009620.
 PA (NICH/) NICHOLS S E.
 PI Nichols SE;
 XX WPI; 2002-414332/44.
 DR Glucosyltransferase B or D protein useful for producing a glucan useful
 PT as substrates for and additions to modified starch and latexes in
 PT paper manufacture, comprises mutations in specific positions -
 XX Claim 36; Page -; 44pp; English.
 XX The invention an isolated protein comprising a glucosyltransferase
 CC (GTF) B polypeptide having changes at position from 1448V, D457N,
 CC D567T, K1014T, D457N/D567T, D457N/D571K, D567T/D571K,
 CC D567T/D571K/K1014T, 1448V/D457N/D567T/D571K/K779Q/K1014T,
 CC Y169A/Y170A/Y171A, and K779Q or a GTF D polypeptide having
 CC changes at positions from T589D, T589E, N471D, N471D/T589D, and
 CC N471D/T589E. Also included are a glucan produced by the GTF mutant,
 CC an isolated polynucleotide which encodes P1 or P2, or its complementary
 CC polynucleotide, a ribonucleic acid sequence encoding the GTF mutant,
 CC an expression cassette comprising the polynucleotide operably linked to a
 CC promoter, a vector comprising the expression cassette, host cell
 CC introduced with the vector, a transgenic plant comprising the
 CC vector, a seed or tuber from the transgenic plant, a paper sizing and/or
 CC coating composition comprising a glucan produced in a plant transformed
 CC with a gene encoding the mutant GTF, wild type or, starch, a latex,
 CC thermoplastic molecule or their combinations or glucan and starch where
 CC the glucan is produced in the amyloplast and/or vacuole or a maize line
 CC deficient in starch biosynthesis, transformed with a gene encoding a
 CC glucosyltransferase B or D enzyme, wild-type or mutant and a paper
 CC comprising the glucan (paper sizing/coating agent). The vector is useful
 CC for producing a glucan in a plant. The method comprises transforming a
 CC plant cell with the vector, growing the plant cell under plant growing
 CC conditions to produce a regenerated plant and inducing expression of the
 CC polynucleotide for a time sufficient to produce the glucan in the
 CC regenerated plant, where the vector contains a transit sequence from
 CC ribulose biphosphate carboxylase small subunit, waxy, brittle-1 and
 CC chlorophyll AB binding protein to produce a transgenic plant, and glucan
 CC is produced in the amyloplast of potato or the vacuole of sugar beet.
 CC Glucans are useful as substrates for and additions to modified starch
 CC and latexes in paper manufacture. Unlike prior art techniques, which
 CC require input materials that produce chemical effluents, paper
 CC manufacture utilizing the glucan produced by GTF, which utilizes
 CC biologically produced input materials, is more cost-effective and
 CC environmentally friendly. Moreover, glucans also exhibit thermoplastic
 CC properties and impart gloss to the paper during coating step.
 CC The present sequence represents a GTFB mutant of the invention.
 CC Note: The present sequence is not shown in the specification but
 CC was created by the indexer using the GTFB sequence appearing as AAU98027

CC and the information in claim 36.
 XX Sequence 1475 AA;
 SQ Query Match 90.8%; Score 108; DB 23; Length 1475;
 Best Local Similarity 85.7%; Pred. No. 2.7e-07;
 Matches 18; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
 QY 1 ANNHVSIEAMSNDNDPTYLHD 21
 DB 481 ANDHLSIEAWSNDNDPTYLHD 501
 RESULT 12
 AAU98038
 ID AAU98038 standard; Protein, 1475 AA.
 AC AAU98038;
 XX 27-AUG-2002 (first entry)
 DT S. mutans GTFB mutant 1448V/D457N/D567T/D571K/K779Q/K1014T.
 XX Glucosyltransferase; GTFB; transgenic plant; paper sizing;
 XX amyloplast; vacuole; glucan; starch; latex; thermoplastic molecule;
 XX Streptococcus mutans.
 OS Synthetic.
 FH Key Location/Qualifiers
 FH Misc-difference 448 /note= "Wild-type Ile substituted by Val"
 FT Misc-difference 457 /note= "Wild-type Asp substituted by Asn"
 FT Misc-difference 567 /note= "Wild-type Asp substituted by Thr"
 FT Misc-difference 571 /note= "Wild-type Asp substituted by Lys"
 FT Misc-difference 779 /note= "Wild-type Lys substituted by Gln"
 FT Misc-difference 1014 /note= "Wild-type Lys substituted by Thr"
 FT US2002031826-A1.
 PN 14-MAR-2002.
 PD 19-DEC-2000; 2000US-0740274.
 XX 11-DEC-1998; 98US-0210361.
 PR 07-JUN-1995; 95US-0478704.
 PR 07-JUN-1995; 95US-0482711.
 PR 07-JUN-1995; 95US-0485243.
 PR 16-JAN-1998; 98US-0007999.
 PR 16-JAN-1998; 98US-0008172.
 PR 20-JAN-1998; 98US-0009620.
 XX (NICH/) NICHOLS S E.
 PI Nichols SE;
 XX WPI; 2002-414332/44.
 DR Glucosyltransferase B or D protein useful for producing a glucan useful
 PT as substrates for and additions to modified starch and latexes in
 PT paper manufacture, comprises mutations in specific positions -
 XX Claim 36; Page -; 44pp; English.
 XX The invention an isolated protein comprising a glucosyltransferase
 CC (GTF) B polypeptide having changes at position from 1448V, D457N,
 CC D567T, K1014T, D457N/D567T, D457N/D571K, D567T/D571K,

D567T/D571K/K1014T, I448V/D457N/D567T/D571K/K779Q/K1014T, Y169A/Y170A/Y171A, and K779Q or a GTF D polypeptide having changes at positions from T589D, T589E, N471D, N471D/T589D and N471D/T589E. Also included are a glucan produced by the GTF mutant, an isolated polynucleotide which encodes P1 or P2, or its complementary polynucleotide, a ribonucleic acid sequence encoding the GTF mutant, an expression cassette comprising the polynucleotide operably linked to a promoter, a vector comprising the expression cassette, host cell introduced with the vector, a transgenic plant comprising the vector, a seed or tuber from the transgenic plant, a paper sizing and/or coating composition comprising a glucan produced in a plant transformed with a gene encoding the mutant GTF, wild type or starch, a latex, thermoplastic molecule or their combinations or glucan and starch where the glucan is produced in the amyloplast and/or vacuole or a maize line deficient in starch biosynthesis, transformed with a gene encoding a glucosyltransferase B or D enzyme, wild-type or mutant and a paper comprising the glucan (paper sizing/coating agent). The vector is useful for producing a glucan in a plant. The method comprises transforming a plant cell with the vector, growing the plant cell under plant growing conditions to produce a regenerated plant and inducing expression of the polynucleotide for a time sufficient to produce the glucan in the regenerated plant, where the vector contains a transit sequence from ribulose biphosphate carboxylase small subunit, waxy, brittle-1 and chlorophyll AB binding protein to produce a transgenic plant, and glucan is produced in the amyloplast of potato or the vacuole of sugar beet. Glucans are useful as substitutes for and additions to modified starch and latexes in paper manufacture. Unlike prior art techniques, which require input materials that produce chemical effluents, paper manufacture utilizing the glucan produced by GTF, which utilizes biologically produced input materials, is more cost-effective and environmentally friendly. Moreover, glucans also exhibit thermoplastic properties and impact gloss to the paper during coating step. The present sequence represents a GTFB mutant of the invention. Note: The present sequence represents a GTFB mutant of the invention. was created by the indexer using the GTFB sequence appearing as AAU98027 and the information in claim 36.

CC Sequence 1475 AA;

CC Query Match 90.8%; Score 108; DB 23; Length 1475;

CC Best Local Similarity 85.7%; Pred. No. 2.7e-07;

CC Matches 18; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

CC 1 ANNHVSIVEAMSDNDPTPLHD 21

CC 481 ANDHLSIEAMSDNDPTPLHD 501

CC RESULT 13

CC AAU98039

CC ID AAU98039 standard; Protein; 1475 AA.

CC XX AAU98039;

CC 27-AUG-2002 (first entry)

CC S. mutans glucosyltransferase GTFB mutant YYY169-171AAA.

CC KM Glucosyltransferase; GTFB; transgenic plant; paper sizing;

CC KM coating composition; glucan; starch; latex; thermoplastic molecule;

CC KM amyloplast; vacuole; paper manufacture; mutant; mutain.

CC XX Streptococcus mutans.

CC OS Synthetic.

CC Key Location/Qualifiers

CC Misc-difference 169..171

CC /note="Wild-type Tyr-Tyr-Tyr substituted by

CC Ala-Ala-Ala"

XX 19-DEC-2000; 2000US-0740274.

XX 11-DEC-1998; 98US-0210361.

XX 07-JUN-1995; 95US-0478704.

XX 07-JUN-1995; 95US-0482711.

XX 07-JUN-1995; 95US-0485243.

XX 16-JAN-1998; 98US-0007999.

XX 16-JAN-1998; 98US-0008172.

XX 20-JAN-1998; 98US-0009620.

XX (NICH/) NICHOLS S. E.

XX Nichols SE;

XX WPI; 2002-414332/44.

XX Glucosyltransferase B or D protein useful for producing a glucan useful as substitutes for and additions to modified starch and latexes in paper manufacture, comprises mutations in specific positions -

XX Claim 36; Page -: 44pp; English.

The invention an isolated protein comprising a glucosyltransferase (GTF) B polypeptide having changes at position from I448V, D457N, D567T, K1014T, D457N/D567T, D457N/D571K, D567T/D571K, D567T/D571K/K1014T, I448V/D457N/D567T/D571K/K779Q/K1014T, Y169A/Y170A/Y171A, and K779Q or a GTF D polypeptide having changes at positions from T589D, T589E, N471D, N471D/T589D, and N471D/T589E. Also included are a glucan produced by the GTF mutant, an isolated polynucleotide which encodes P1 or P2, or its complementary polynucleotide, a ribonucleic acid sequence encoding the GTF mutant, an expression cassette comprising the polynucleotide operably linked to a promoter, a vector comprising the expression cassette, host cell introduced with the vector, a transgenic plant comprising the vector, a seed or tuber from the transgenic plant, a paper sizing and/or coating composition comprising a glucan produced in a plant transformed with a gene encoding the mutant GTF, wild type or starch, a latex, thermoplastic molecule or their combinations or glucan and starch where the glucan is produced in the amyloplast and/or vacuole or a maize line deficient in starch biosynthesis, transformed with a gene encoding a glucosyltransferase B or D enzyme, wild-type or mutant and a paper comprising the glucan (paper sizing/coating agent). The vector is useful for producing a glucan in a plant. The method comprises transforming a plant cell with the vector, growing the plant cell under plant growing conditions to produce a regenerated plant and inducing expression of the polynucleotide for a time sufficient to produce the glucan in the regenerated plant, where the vector contains a transit sequence from ribulose biphosphate carboxylase small subunit, waxy, brittle-1 and chlorophyll AB binding protein to produce a transgenic plant, and glucan is produced in the amyloplast of potato or the vacuole of sugar beet. Glucans are useful as substitutes for and additions to modified starch and latexes in paper manufacture. Unlike prior art techniques, which require input materials that produce chemical effluents, paper manufacture utilizing the glucan produced by GTF, which utilizes biologically produced input materials, is more cost-effective and environmentally friendly. Moreover, glucans also exhibit thermoplastic properties and impact gloss to the paper during coating step. The present sequence represents a GTFB mutant of the invention. Note: The present sequence is not shown in the specification but was created by the indexer using the GTFB sequence appearing as AAU98027 and the information in claim 36.

CC Sequence 1475 AA;

CC Query Match 90.8%; Score 108; DB 23; Length 1475;

CC Best Local Similarity 85.7%; Pred. No. 2.7e-07;

CC Matches 18; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

CC 1 ANNHVSIVEAMSDNDPTPLHD 21

CC 481 ANDHLSIEAMSDNDPTPLHD 501

ID	AAU98040	standard; Protein; 1475 AA.
XX	AAU98040;	
XX	27-AUG-2002 (first entry)	
XX	S. mutans glucosyltransferase GTFB mutant K779Q.	
XX	Glucosyltransferase; GTFB; transgenic plant; paper sizing;	
KW	coating composition; glucan; starch; latex; thermoplastic molecule;	
KW	amyloplast; vacuole; paper manufacture; mutant; mutain.	
OS	Streptococcus mutans.	
OS	Synthetic.	
XX	Key	Location/Qualifiers
XX	Misc-difference 779	/note= "Wild-type Lys substituted by Gln"
XX	US2002031826-A1.	
XX	14-MAR-2002.	
XX	19-DEC-2000; 2000US-0740274.	
XX	11-DEC-1998; 98US-0210361.	
XX	07-JUN-1995; 95US-0478704.	
XX	07-JUN-1995; 95US-0482711.	
XX	07-JUN-1995; 95US-0485243.	
XX	16-JAN-1998; 98US-0007999.	
XX	16-JAN-1998; 98US-0008172.	
XX	20-JAN-1998; 98US-0009620.	
XX	(NICH/) NICHOLS S E.	
XX	Nichols SE;	
XX	WPI; 2002-414332/44.	
XX	Glucosyltransferase B or D protein useful for producing a glucan useful	
XX	as substitutes for and additions to modified starch and latexes in	
XX	paper manufacture, comprises mutations in specific positions -	
XX	Claim 36; Page -; 44pp; English.	
XX	The invention an isolated protein comprising a glucosyltransferase	
XX	(GTF) B polypeptide having changes at position from 1448V, D457N,	
XX	D567T, K1014T, D457N/D567T, D457N/D571K, D567T/D571K,	
XX	D567T/D571K/K1014T, 1448V/D457N/D567T/D571K/K779Q/K1014T,	
XX	Y169A/Y170A/Y171A, and K779Q or a GTF D polypeptide having	
XX	changes at positions from 1589D, 1589E, N471D, N471D/1589D, and	
XX	N471D/1589E. Also included are a glucan produced by the GTF mutant,	
XX	an isolated polynucleotide which encodes p1 or p2, or its complementary	
XX	polynucleotide, a ribonucleic acid sequence encoding the GTF mutant,	
XX	an expression cassette comprising the polynucleotide operably linked to a	
XX	promoter, a vector comprising the expression cassette, host cell	
XX	introduced with the vector, a transgenic plant comprising the	
XX	vector, a seed or tuber from the transgenic plant, a paper sizing and/or	
XX	coating composition comprising a glucan produced in a plant transformed	
XX	with a gene encoding the mutant GTF, wild type or, starch, a latex,	
XX	thermoplastic molecule or their combinations or glucan and starch where	
XX	the glucan is produced in the amyloplast and/or vacuole or a maize line	
XX	deficient in starch biosynthesis, transformed with a gene encoding a	
XX	glucosyltransferase B or D enzyme, wild-type or mutant and a paper	
XX	comprising the glucan (paper sizing/coating agent). The vector is useful	
XX	for producing a glucan in a plant. The method comprises transforming a	
XX	plant cell with the vector, growing the plant cell under plant growing	
XX	conditions to produce a regenerated plant and inducing expression of the	
XX	polynucleotide for a time sufficient to produce the glucan in the	
XX	regenerated plant, where the vector contains a transit sequence from	

CC ribulose biphosphate carboxylase small subunit, waxy, brittle-1 and glu
CC chlorophyll AB binding protein to produce a transgenic plant, and glucan
CC is produced in the amyloplast of potato or the vacuole of sugar beet.
CC Glucans are useful as substitutes for and additions to modified starch
CC and lakeas in paper manufacture. Unlike prior art techniques, which
CC require input materials that produce chemical effluents, paper
CC manufacture utilizing the glucan produced by GTF, which utilizes
CC biologically produced input materials, is more cost-effective and
CC environmentally friendly. Moreover, glucans also exhibit thermoplastic
CC properties and impart gloss to the paper during coating step.
CC The present sequence represents a GTFB mutant of the invention.
CC Note: The present sequence is not shown in the specification but
CC was created by the indexer using the GTFB sequence appearing as AA098027
CC and the information in claim 36.

XX SQ Sequence 1475 AA;

QY 1 ANNEHSIVEAWSNDPTPYLHD 21
DB 481 ANDEHLSLEAWSNDPTPYLHD 501

Matches 18; Conservative 3; Mismatches 0; Indels 0; Gaps 0

Query Match 90.8%; Score 108; DB 23; Length 1475;
Best Local Similarity 85.7%; Pred.No. 2.7e-07;
Matches 18; Conservative 3; Mismatches 0; Indels 0; Gaps 0

RESULT 15
AAU79284
ID AAU79284 standard; Protein; 1476 AA.

AC AAU79284;
XX
XX
DT 13-AUG-2002 (first entry)
DE Streptococcus mutans monoclonal antibody-related protein #1.
KM Streptococcus mutans monoclonal antibody-related protein #1.
KW Antibody; dental caries; water insoluble glucan synthetase;
XX anti-carries; glucosyl transferase-B; immunotherapy.
XX
XX Streptococcus mutans.
OS
XX JP2002114709-A.
PN
XX
PD 16-APR-2002.
XX
PF 04-OCT-2000; 2000JP-0304889.
XX
PR 04-OCT-2000; 2000JP-0304889.
XX
PA (UNIT-) UNIT NIPPON.
XX
DR WPI, 2002-448101/48.
XX
XX Anti-carries agent composed of a monoclonal antibody against an
PT inhibitory enzyme against water insoluble glucan synthetase of glucosyl
PT transferase-B (GTF-B) of Streptococcus mutans -
PT
PS Claim 3; Page 13-16; 28pp; Japanese.
XX
XX The invention relates to a monoclonal antibody against dental caries and
CC an anti-carries agent composed of a monoclonal antibody produced by
CC Streptococcus mutans, particularly mouse-hybridoma MHP126 (FERM P-17566)
CC or mouse-Hybridoma MHP136 (FERM P-17567), against an enzyme having
CC inhibitive activity against water insoluble glucan synthetase of glucosyl
CC transferase-B. The monoclonal antibody specifically inhibits water
CC insoluble glucan synthetase of Streptococcus mutans produced glucosyl
CC transferase-B and is used in the immunotherapy of dental caries. This
CC sequence represents a Streptococcus mutans monoclonal antibody-related
CC protein.
XX
XX SQ Sequence 1476 AA;

Query Match 90.8%; Score 108; DB 23; Length 1476;

Thu Nov 13 12:12:11 2003

us-09-290-049a-14.rag

Page 11

Best Local Similarity 85.7%; Pred. No. 2.7e-07;
Matches 18; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

```
QY      1 ANNHVSIVEAWSNDNTPYIHD 21
      ||:|:|:|:|:|:|:|:|:|
Db     481 ANDHLSILEAWSNDNTPYIHD 501
```

Search completed: November 13, 2003, 09:38:28
Job time : 39.5166 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 13, 2003, 09:31:40 ; Search time 13.0379 Seconds
(without alignments)
154.898 Million cell updates/sec

Title: US-09-290-049a-14
Perfect score: 119
Sequence: 1 ANNHVSIVEAMSDNDTPYLHD 21

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: PIR 76: *
2: PIR1: *
3: PIR2: *
4: PIR3: *
5: PIR4: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	119	100.0	1592	2 A38175	glucosyltransferase
2	108	90.8	1475	2 B31135	glb protein precursor
3	99	83.2	1375	2 J70345	dextranase (EC
4	74	62.2	1431	2 A45866	dextranase (EC
5	68	57.1	1508	2 T31098	probable dextran
6	67	56.3	1365	2 A41483	glucosyltransferase
7	61	51.3	1449	2 T30857	glucosyltransferase
8	61	51.3	1449	2 T30552	glucosyltransferase
9	57	47.9	1577	2 T30858	glucosyltransferase
10	56	47.1	1518	2 A44811	glucosyltransferase
11	55	46.2	1599	2 S22737	glucosyltransferase
12	53	44.5	175	2 C86205	hypothetical prote
13	49	41.2	336	1 CQBR40	BGRF2 protein - hu
14	49	41.2	524	2 D82220	conserved hypothet
15	48	40.3	331	2 B48445	glyceroldehyde-3-p
16	46	38.7	313	2 S59448	hypothetical prote
17	46	38.7	490	2 H70538	probable ppdk prot
18	46	38.7	601	2 E87028	pyruvate, phosphat
19	45	38.2	418	2 D90506	4-aminobutyrate am
20	45	37.8	337	2 C64233	glyceroldehyde-3-p
21	45	37.8	347	2 T48610	hypothetical prote
22	45	37.8	525	2 T40088	RhoGEF domain cont
23	45	37.8	947	2 B63662	hypothetical prote
24	45	37.8	1384	2 T02748	hypothetical prote
25	45	37.8	4848	2 T30289	pristinamycin I sy
26	44.5	37.4	835	2 B64689	site-specific DNA-
27	44	37.0	419	2 A59414	metalloprotease
28	44	37.0	741	2 JC5142	X-Pro dipeptidyl-p
29	44	37.0	1122	2 S64443	probable membrane

30	44	37.0	1131	2 T16217	hypothetical prote
31	44	37.0	1313	2 T29027	hypothetical prote
32	43.5	36.6	491	1 D64947	glucose-6-phosphat
33	43.5	36.6	491	2 B90949	glucose-6-phosphat
34	43.5	36.6	491	2 F85797	glucose-6-phosphat
35	43.5	36.6	491	2 AB0742	glucose-6-phosphat
36	43	36.1	78	2 S76593	hypothetical prote
37	43	36.1	194	2 H72037	conserved hypothet
38	43	36.1	194	2 C86886	CT647 hypothetical
39	43	36.1	275	2 S55978	hypothetical prote
40	43	36.1	284	2 UC7148	heat-shock sigma f
41	43	36.1	331	1 DBUTGC	glyceroldehyde-3-p
42	43	36.1	346	2 E70715	hypothetical prote
43	43	36.1	351	2 D89991	hypothetical prote
44	43	36.1	396	2 T35024	probable glutathio
45	43	36.1	441	2 T36737	probable homogenit

ALIGNMENTS

RESULT 1

A38175 glucosyltransferase precursor - Streptococcus sobrinus
C/Species: Streptococcus sobrinus
C/Date: 28-Aug-1992 #sequence_revision 28-Aug-1992 #text_change 15-Oct-1999
C/Accession: A38175
R/Abn, H.; Mtsamura, T.; Kodama, T.; Ohta, H.; Fukui, K.; Kato, K.; Kagawa, H.
U. Bacteriol. 173, 989-996, 1991
A/Title: Peptide sequences for sucrose splitting and glucan binding within Streptococcus
A/Reference number: A38175; MUID:91123227; PMID:1704006
A/Accession: A38175
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-1592 <ABO>
A/Cross-references: GB:D90213; NID:G217032; PIDN:BA14241.1; PID:d1014946; PID:G217033
C/Superfamily: cpl repeat homology
F/1093-1113/Domain: cpl repeat homology <CP1>
F/1222-1241/Domain: cpl repeat homology <CP2>
F/1287-1306/Domain: cpl repeat homology <CP3>
F/1330-1351/Domain: cpl repeat homology <CP4>
F/1352-1371/Domain: cpl repeat homology <CP5>
F/1402-1420/Domain: cpl repeat homology <CP6>
F/1465-1484/Domain: cpl repeat homology <CP7>
F/1513-1532/Domain: cpl repeat homology <CP8>
Query Match 100.0%; Score 119; DB 2; Length 1592;
Best local similarity 100.0%; Pred. No. 5.6e-10;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 ANNHVSIVEAMSDNDTPYLHD 21
477 ANNHVSIVEAMSDNDTPYLHD 497

RESULT 2
B31135 glb protein precursor - Streptococcus mutans
C/Species: Streptococcus mutans
C/Date: 23-Oct-1990 #sequence_revision 23-Oct-1990 #text_change 15-Oct-1999
C/Accession: B31135; A33128
R/Shiroza, T.; Ueda, S.; Kuramitsu, H. K.
J. Bacteriol. 169, 4263-4270, 1987
A/Title: Sequence analysis of the glbF gene from Streptococcus mutans.
A/Reference number: A33135; MUID:87308013; PMID:3040685
A/Accession: B31135
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-1475 <SHI>
A/Cross-references: GB:M17361; NID:G153639; PIDN:AAA8588.1; PID:G153640
R/Shiroza, T.; Ueda, S.; Kuramitsu, H. K.
Submitted to the Protein Sequence Database, September 1990
A/Reference number: A33128

```

A/Accession: A33128
A/Status: preliminary; not compared with conceptual translation
A/Molecule type: DNA
A/Residues: 1-171,173-641,'N',643-1475 <SH2>
A/Experimental source: strain GS-5
A/Superfamily: cpl repeat homology
F/1096-1115/Domain: cpl repeat homology <CP1>
F/1224-1243/Domain: cpl repeat homology <CP2>
F/1289-1308/Domain: cpl repeat homology <CP3>
F/1354-1373/Domain: cpl repeat homology <CP4>
F/1419-1438/Domain: cpl repeat homology <CP5>

Query Match          90.8%; Score 108; DB 2; Length 1475;
Best Local Similarity 85.7%; Pred. No. 2.6e-08;
Matches 18; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ANNHVSIVEAWSNDNTPYLHD 21
   |||:|||||
Db 481 ANDHLSILEAWSNDNTPYLHD 501

RESULT 3
UT0345
dextranucrase (EC 2.4.1.5) precursor - Streptococcus mutans (strain GS-5)
N/Alternate names: sucrose 6-glucosyltransferase
C/Species: Streptococcus mutans
C/Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 24-Sep-1993
C/Accession: JT0345; C31135
R/Ueda, S.; Shiroza, T.; Kuramitsu, H.K.
Gene 69, 101-109, 1988
A/Title: Sequence analysis of the gtfC gene from Streptococcus mutans GS-5.
A/Reference number: JT0345; PMID:89137980; PMID:2576010
A/Accession: JT0345
A/Molecule type: DNA
A/Residues: 1-1375 <UED>
A/Experimental source: GS-5
R/Shiroza, T.; Ueda, S.; Kuramitsu, H.K.
J. Bacteriol. 169, 4263-4270, 1987
A/Title: Sequence analysis of the gtfB gene from Streptococcus mutans.
A/Reference number: A33135; PMID:87308013; PMID:3040685
A/Accession: C31135
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-349 <SH1>
A/Cross-references: GB:M17361
A/Genetics:
A/Gene: gtfC
C/Function:
A/Description: catalyzes the synthesis of both water-soluble and water-insoluble glucans
C/Superfamily: cpl repeat homology
C/Keywords: duplication; glycosyltransferase; hexosyltransferase
F/1-34/Domain: signal sequence #status predicted <SIG>
F/35-1375/Product: glucosyltransferase #status predicted <MAT>
F/1126-1145/Domain: cpl repeat homology <CP1>
F/1253-1272/Domain: cpl repeat homology <CP2>
F/1318-1337/Domain: cpl repeat homology <CP3>

Query Match          83.2%; Score 99; DB 2; Length 1375;
Best Local Similarity 81.0%; Pred. No. 6.1e-07;
Matches 17; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ANNHVSIVEAWSNDNTPYLHD 21
   |||:|||||
Db 507 ANDHLSILEAWSNDNTPYLHD 527

RESULT 4
A45866
dextranucrase (EC 2.4.1.5) precursor - Streptococcus mutans
C/Species: Streptococcus mutans
C/Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 07-May-1999
C/Accession: A45866
R/Honda, O.; Kato, C.; Kuramitsu, H.K.

```

```

J. Gen. Microbiol. 136, 2099-2105, 1990
A/Title: Nucleotide sequence of the Streptococcus mutans gtfD gene encoding the glucosyl
A/Reference number: A45866; PMID:91100958; PMID:2148600
A/Accession: A45866
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-1431 <HON>
A/Cross-references: GB:M29296
C/Superfamily: cpl repeat homology
C/Keywords: glycosyltransferase; hexosyltransferase
F/181-201/Domain: cpl repeat homology <CP1>
F/1127-1146/Domain: cpl repeat homology <CP2>
F/1192-1211/Domain: cpl repeat homology <CP3>
F/1257-1276/Domain: cpl repeat homology <CP4>
F/1277-1297/Domain: cpl repeat homology <CP5>
F/1321-1340/Domain: cpl repeat homology <CP6>
F/1341-1361/Domain: cpl repeat homology <CP6>
F/1385-1404/Domain: cpl repeat homology <CP7>

Query Match          62.2%; Score 74; DB 2; Length 1431;
Best Local Similarity 66.7%; Pred. No. 0.005;
Matches 14; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Qy 1 ANNHVSIVEAWSNDNTPYLHD 21
   |||:|||||
Db 495 AINHLSTLEAWSNDNTPYKND 515

RESULT 5
T31098
probable dextranucrase (EC 2.4.1.5), extracellular - Leuconostoc mesenteroides
C/Species: Leuconostoc mesenteroides
C/Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 11-May-2000
C/Accession: T31098
R/Monchois, V.; Renaud-Simeon, M.; Monsan, P.; Willemot, R.M.
FEMS Microbiol. Lett. 159, 307-315, 1998
A/Title: Cloning and sequencing of a gene coding for an extracellular dextranucrase (DS
A/Reference number: Z20981; PMID:98164374; PMID:9503626
A/Accession: T31098
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-1508 <MON>
A/Cross-references: EMBL:AF030129; NID:92766611; PID:92766612; PIND:AA895453.1
A/Experimental source: strain NRRL B-1299
A/Genetics:
A/Gene: dsrB
C/Function:
A/Description: produces dextran composed only of alpha(1-6) glucosidic bonds
C/Keywords: glycosyltransferase; hexosyltransferase

Query Match          57.1%; Score 68; DB 2; Length 1508;
Best Local Similarity 57.1%; Pred. No. 0.046;
Matches 12; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

Qy 1 ANNHVSIVEAWSNDNTPYLHD 21
   |||:|||||
Db 563 ANQHLSLEAWSNDNTPYKND 583

RESULT 6
A41483
glucosyltransferase (EC 2.4.1.-) gtfS precursor - Streptococcus sobrinus
C/Species: Streptococcus sobrinus
C/Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 15-Oct-1999
C/Accession: A41483
R/Gilmore, K.S.; Russell, R.R.B.; Ferretti, J.J.
Infect. Immun. 58, 2452-2458, 1990
A/Title: Analysis of the Streptococcus downei gtfS gene, which specifies a glucosyltrans
A/Reference number: A41483; PMID:9031665; PMID:2142479
A/Accession: A41483
A/Molecule type: DNA
A/Residues: 1-1365 <GLI>
A/Cross-references: GB:M30943; NID:g153652; PIND:AAA6898.1; PID:g153653

```

```
C:/Genetics:
A:/Gene: gtfS
C:/Superfamily: cpl repeat homology
C:/Keywords: glycosyltransferase; hexosyltransferase

Query Match      56.3%; Score 67; DB 2; Length 1365;
Best Local Similarity 57.1%; Pred. No. 0.058;
Matches 12; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY      1 ANNHVSIVEAWSNDNTPYLHD 21
       |:|::|||::| |::|
Db      467 AIKHLSTLEAWSGNDNYVD 487

RESULT 7
T30857
glucosyltransferase - Streptococcus salivarius
C:/Species: Streptococcus salivarius
C:/Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 22-Oct-1999
C:/Accession: T30857
R:/Simson, C.L.; Giffard, P.M.; Jacques, N.A.
Infect. Immun. 63, 609-621, 1995
A:/Title: Streptococcus salivarius ATCC 25975 possesses at least two genes coding for pro
A:/Reference number: Z20909; PMID:95122197; PMID:7822030
A:/Accession: T30857
A:/Status: preliminary; translated from GB/EMBL/DDBJ
A:/Molecule type: DNA
A:/Residues: 1-1449 <SIM>
A:/Cross-references: EMBL:L35495; NID:g662378; PID:g662379; PIDN:AAC41412.1
C:/Genetics:
A:/Gene: gtlF

Query Match      51.3%; Score 61; DB 2; Length 1449;
Best Local Similarity 57.1%; Pred. No. 0.54;
Matches 12; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY      1 ANNVSVTEAWSDNDTPYLHD 21
       |:|::|||::| |::|
Db      536 AIKHLSILEAWSHNDAYINED 556

RESULT 8
T30552
glucosyltransferase N - Streptococcus salivarius (fragment)
C:/Species: Streptococcus salivarius
C:/Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 22-Oct-1999
C:/Accession: T30552
R:/Jaffe, R.I.
submitted to the EMBL Data Library, February 1998
A:/Description: Streptococcus salivarius V477 gtfN.
A:/Reference number: Z20854
A:/Accession: T30552
A:/Status: preliminary; translated from GB/EMBL/DDBJ
A:/Molecule type: DNA
A:/Residues: 1-1449 <JAF>
A:/Cross-references: EMBL:AF049609; NID:g2935545; PID:g2935546; PIDN:AAC05156.1
C:/Genetics:
A:/Gene: gtfN

Query Match      51.3%; Score 61; DB 2; Length 1449;
Best Local Similarity 57.1%; Pred. No. 0.54;
Matches 12; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY      1 ANNHVSIVEAWSNDNTPYLHD 21
       |:|::|||::| |::|
Db      536 AIKHLSILEAWSHNDAYINED 556

RESULT 9
T30858
glucosyltransferase - Streptococcus salivarius
C:/Species: Streptococcus salivarius
C:/Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 22-Oct-1999
```

```
C:\jackson> J:\Giffard_P.M.; Jacques_N.A.  
Infect. Immun. 63, 609-621, 1995  
A>Title: Streptococcus salivarius ATCC 25975 possesses at least two genes coding for pil  
A.Reference number: Z20909; MUID:J95122197; PMID:7822030  
A.Accession: F30858  
A>Status: preliminary; translated from GB/EMBL/DDBJ  
A:Molecule type: DNA  
A:Residues: 1-1577 <STM>  
A:Cross-references: EMBL:L35928; NID:g662380; PID:g662381; PIND:AAC41413.1  
C:Genetics:  
A:gene: gtfm
```

```
Query Match      47.9%; Score 57; DB 2; Length 1577;  
Best Local Similarity    61.1%; Pred.No. 2.5;  
Matches          11; Conservative     2; Mismatches       5; Indels         0; Gaps        0;
```

```
OY           4 HVSIVAMSDNDTPYLHD 21  
|:-|-||| |  
Db            591 HTSLIFAMSYNHQMVKD   608
```

```
RESULT 10  
A44811  
Glucosyltransferase (EC 2.4.1.-) I - Streptococcus salivarius  
C.Species: Streptococcus salivarius  
C.Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 15-Oct-1999  
R.Citation: A44811; S22726; S28809  
R.Giffard_P.M.; Simpson_C.L.; Milward_C.P.; Jacques_N.A.  
J.Gen.Microbiol. 137, 2577-2593, 1991  
A.Title: Molecular characterization of a cluster of at least two glucosyltransferase gene  
A.Reference number: A44811; MUID:92148377; PMID:1858391  
A.Accession: A44811  
A:Molecule type: DNA  
A.Residues: 1-1518 <GLF>  
A.Cross-references: EMBL:Z11873; NID:g47526; PIND:CAA77900.1; PID:g47527  
A>Note: sequence extracted from NCBI backbone (NCBIN:81050, NCBI:81052)  
C:Genetics:  
A:Gene: glfU  
C:Superfamily: cpl repeat homology  
C:Keywords: glycosyltransferase; hexosyltransferase  
P.I307-I326/Domain: cpl repeat homology <CP4>
```

```
Query Match      47.1%; Score 56; DB 2; Length 1518;  
Best Local Similarity    60.0%; Pred.No. 3.4;  
Matches           9; Conservative     3; Mismatches       3; Indels         0; Gaps        0;
```

```
OY           4 HVSIVEAWSMDNPTPY 18  
|:-|-||| |  
Db            504 HTSVLEAFMSLNDRHY 518
```

```
RESULT 11  
S22737  
glucosyltransferase (EC 2.4.1.-) S - Streptococcus salivarius  
C.Species: Streptococcus salivarius  
C.Date: 16-Sep-1992 #sequence_revision 16-Sep-1992 #text_change 21-Jan-2000  
R.Citation: S22737; S28810; B44811; S22727  
R.Jacques_N.  
submitted to the EMBL Data Library, March 1992  
A.Reference number: S22726  
A.Accession: S22737  
A:Molecule type: DNA  
A:Residues: 1-1599 <JAC>  
A:Cross-references: EMBL:Z11872; NID:g47530; PIND:CAA77898.1; PID:g47531  
A.Experimental source: AFCC 25975  
R.Giffard_P.M.; Simpson_C.L.; Milward_C.P.; Jacques_N.A.  
J.Gen.Microbiol. 137, 2577-2593, 1991  
A>Title: Molecular characterization of a cluster of at least two glucosyltransferase gen  
A.Reference number: A44811; MUID:92148377; PMID:1858391  
A.Accession: S28810  
A:Molecule type: DNA  
A:Residues: 1-51 <GIT>
```


A:Cross-references: EMBL:Z11873

C:Genetics:

A:Gene: gtlK

C:Superfamily: cpl repeat homology

C:Keywords: glycosyltransferase; hexosyltransferase

F:1456-1475/Domain: cpl repeat homology <CPR>

Query Match 46.2%; Score 55; DB 2; Length 1599;
Best Local Similarity 50.0%; Pred. No. 5.2;
Matches 9; Conservative 6; Mismatches 3; Indels 0; Gaps 0;

QY 4 NNVSIYEAMSDNDTPYLHD 21
DB 494 NISILEAMSHNDPYVNE 511

RESULT 12

C66205
hypothetical protein [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C>Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001

C:Accession: C66205

R:Thelogs, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,

Chin, C.W.; Chung, M.K.; Com, L.; Conway, A.B.; Creasy, T.H.; Dewar, K.;

ansen, N.F.; Hughes, B.; Huizar, L.

Nature 408, 816-820, 2000

A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.

C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luos, J.S.; Malet, R.; Marziani,

Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A:Authors: Salzbarg, S.L.; Schwartz, J.R.; Shim, P.; Southwick, A.M.; Sun, H.; Tallon,

ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.M.

A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A:Reference number: A86141; MUID:21016719; PMID:11130712

C:Accession: C66205

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-175 <STO>

A:Cross-references: GB:AE005172; NID:g9954041; PIDN:AAF2215.1; GSPDB:GN00141

C:Genetics:

A:Map position: 1

Query Match 44.5%; Score 53; DB 2; Length 175;
Best Local Similarity 42.1%; Pred. No. 0.78;
Matches 8; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

QY 3 NNVSIYEAMSDNDTPYLHD 21
DB 109 NHOEVIDAMSDHOKPLMTD 127

RESULT 13

Q6840
BgLF2 protein - human herpesvirus 4 (strain B95-8)

C:Species: human herpesvirus 4, Epstein-Barr virus

C>Date: 25-Feb-1985 #sequence_revision 25-Feb-1985 #text_change 16-Jul-1999

C:Accession: C43044; J01381; A03784; A03794; S33036

R:Bankier, A.T.; Deininger, P.L.; Farrell, P.D.; Barrell, B.G.

Mol. Biol. Med. 1, 21-45, 1983

A:Title: Sequence analysis of the 17,166 bp EcoRI fragment C of B95-8 Epstein-Barr virus

A:Reference number: A93065; MUID:85035713; PMID:6092825

C:Accession: C43044

A:Molecule type: DNA

A:Residues: 1-336 <HAN>

A:Cross-references: EMBL:V01555; NID:g59074; PIDN:CAA24831.1; PID:g1334895

R:Bankier, A.T.; Bigngin, M.D.; Deininger, P.L.; Farrell, P.D.; Gibson, T.J.; H

Nature 310, 207-211, 1984

A:Title: DNA sequence and expression of the B95-8 Epstein-Barr virus genome.

A:Reference number: A03794; MUID:84270667; PMID:6087149

C:Contents: annotation; protein coding region

A:Note: neither amino acid nor nucleotide sequence is given

R:Chen, M.R.; Heu, T.Y.; Lin, S.W.; Yang, C.S.

J. Gen. Virol. 72, 3047-3055, 1991

A:Title: Cloning and characterization of cDNA clones corresponding to transcripts from

A:Reference number: J01381; MUID:92113548; PMID:1662696

C:Accession: J01381

A:Molecule type: mRNA

A:Residues: 1-336 <CHE>

A:Cross-references: GB:S7732; NID:g243314; PIDN:AA821113.1; PID:g243315

C:Superfamily: Epstein-Barr virus BgLF2 protein

Query Match 41.2%; Score 49; DB 1; Length 336;
Best Local Similarity 41.2%; Pred. No. 7.1;
Matches 7; Conservative 6; Mismatches 4; Indels 0; Gaps 0;

QY 1 NNVSIYEAMSDNDTPYLHD 17
DB 216 AGAHVNIIRGWTEDDSP 232

RESULT 14

D82220
conserved hypothetical protein VC1268 [imported] - Vibrio cholerae (strain N16961 serogr

C:Species: Vibrio cholerae

C>Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Feb-2001

C:Accession: D82220

R:Heidelberg, J.F.; Eissen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwin, M.L.; Dodson, R.J.;

Charlson, D.; Ermolaeva, M.D.; Vamathevan, J.; Basu, S.; Qin, H.; Dragoi, I.; Sellers, P

1, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.

Nature 406, 477-483, 2000

A:Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.

A:Reference number: A82035; MUID:20406833; PMID:10952301

C:Accession: D82220

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-524 <HEI>

A:Cross-references: GB:AE004206; GB:AE003852; NID:g9655749; PIDN:AAF94427.1; GSPDB:GN001

A:Experimental source: serogroup O1, strain N16961; biotype El Tor

C:Genetics:

A:Gene: VC1268

A:Map position: 1

Query Match 41.2%; Score 49; DB 2; Length 524;
Best Local Similarity 43.5%; Pred. No. 12;
Matches 10; Conservative 4; Mismatches 7; Indels 2; Gaps 1;

QY 1 NNVSIYEAMSDNDTPYLHD 21
DB 365 ANHHWEIIRGWSDEVIDPALID 387

RESULT 15

B48445
glyceraldehyde-3-phosphate dehydrogenase (phosphorylating) (EC 1.2.1.12) - Leishmania me

C:Species: Leishmania mexicana

C>Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 03-Jun-2002

C:Accession: B48445; S25142

R:Hannett, V.; Blauw, M.; Kohl, L.; Allert, S.; Oppendoes, F.R.; Michels, P.A.M.

Mol. Biochem. Parasitol. 55, 115-126, 1992

A:Title: Molecular analysis of the cytosolic and glycosomal glyceraldehyde-3-phosphate d

A:Reference number: A48445; MUID:93063042; PMID:1435864

C:Accession: B48445

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-331 <HAN>

A:Cross-references: EMBL:X65220; NID:g9552; PIDN:CAA46323.1; PID:g9553

C:Superfamily: glyceraldehyde-3-phosphate dehydrogenase

C:Keywords: oxidoreductase

Query Match 40.3%; Score 48; DB 2; Length 331;
Best Local Similarity 42.1%; Pred. No. 9.9;
Matches 8; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

QY 2 NNVSIYEAMSDNDTPYLHD 20
DB 301 NHHFVXLVSWYNGEYSH 319

Thu Nov 13 12:12:12 2003

us-09-290-049a-14.rpr

Page 5

Search completed: November 13, 2003, 09:50:28
Job time : 14.0379 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 13, 2003, 09:31:40 ; Search time 7.56398 Seconds
(without alignments)
130.561 Million cell updates/sec

Title: US-09-290-049A-14

Perfect score: 119
Sequence: 1 ANNHVSIVEAMSDNDTPYLHD 21

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	119	100.0	1592	1	GTF2_STRDO
2	119	100.0	1592	1	GTF1_STRDO
3	108	90.8	1476	1	GTFB_STRMU
4	99	83.2	1455	1	GTFB_STRMU
5	74	62.2	1462	1	GTFD_STRMU
6	67	56.3	1365	1	GTFB_STRDO
7	49	41.2	336	1	UUL6_EBV
8	48	40.3	330	1	G3PC_LETME
9	46	38.7	313	1	YMS8_YEAST
10	45	37.8	337	1	G3P_MYCGE
11	44	37.0	569	1	SLIF_MOUSE
12	44	37.0	1122	1	YG3C_YEAST
13	44	37.0	1131	1	YPO4_CARBL
14	44	37.0	3358	1	PGCV_MOUSE
15	43.5	36.6	491	1	G6PD_ECOLI
16	43	36.1	275	1	SC35_YEAST
17	43	36.1	330	1	G3PC_TRYBB
18	43	36.1	346	1	Y943_MYCUC
19	43	36.1	441	1	HGD_STRCO
20	43	36.1	603	1	POPO_HUMAN
21	43	36.1	899	1	SUHM_DROYI
22	42.5	35.7	470	1	AMP2_STRCO
23	42.5	35.7	491	1	G6PD_ERWCH
24	42	35.3	270	1	PSF1_DROME
25	42	35.3	279	1	G316_YEAST
26	42	35.3	334	1	G3P_BACST
27	42	35.3	335	1	G3P_BACCO
28	42	35.3	425	1	HGD_CAUCR
29	42	35.3	432	1	HGD_PSEAR
30	42	35.3	438	1	GATD_PYRPU
31	42	35.3	521	1	SYM_UREPA
32	42	35.3	533	1	YE09_MYCPN
33	42	35.3	775	1	AD28_HUMAN

ALIGNMENTS

RESULT 1	ID	STANDARD	PRT	1592 AA	DESCRIPTION
GTF2_STRDO	AC	GTF2_STRDO			
DT	01-AUG-1992 (Rel. 23, Created)				
DT	01-AUG-1992 (Rel. 23, Last sequence update)				
DT	15-SEP-2003 (Rel. 42, Last annotation update)				
DE	Glucosyltransferase-I precursor (EC 2.4.1.5) (GTF-I) (Dextranucrase)				
DE	(sucrose 6-glucosyltransferase).				
OS	Streptococcus downei (Streptococcus sobrinus).				
OC	Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;				
OC	Streptococcus.				
OX	NCBI_Taxid=1317;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=6715 / Serotype G;				
RX	MEDLINE=9113227; PubMed=1704006;				
RA	Abo H., Matsumura T., Kodama T., Ohta H., Fukui K., Kato K.,				
RA	Kagawa H.;				
RT	"peptide sequences for sucrose splitting and glucan binding within				
RT	Streptococcus sobrinus glucosyltransferase (water-insoluble glucan				
RT	synthetase).";				
RL	J. Bacteriol. 173:989-996(1991).				
CC	-1- FUNCTION: PRODUCTION OF EXTRACELLULAR GLUCANS, THAT ARE THOUGHT				
CC	TO PLAY A KEY ROLE IN THE DEVELOPMENT OF THE DENTAL PLAQUE BECAUSE				
CC	OF THEIR ABILITY TO ADHERE TO SMOOTH SURFACES AND MEDIATE THE				
CC	AGGREGATION OF BACTERIAL CELLS AND FOOD DEBRIS.				
CC	-1- CATALYTIC ACTIVITY: Sucrose + ((1,6)-alpha-D-glucosyl)(N) = D-				
CC	fructose + ((1,6)-alpha-D-glucosyl)(N+1).				
CC	-1- SUBCELLULAR LOCATION: Secreted.				
CC	-1- DISEASE: DENTAL CARIES.				
CC	-1- MISCELLANEOUS: GTF-I SYNTHESIZES WATER-INSOLUBLE GLUCANS (ALPHA				
CC	1,3-LINKED GLUCOSE AND SOME 1,6-LINKAGES). GTF-S SYNTHESIZES				
CC	WATER-SOLUBLE GLUCANS (ALPHA 1,6-GLUCOSE). GTF-SI SYNTHESIZES BOTH				
CC	FORMS OF GLUCANS.				
CC	-1- SIMILARITY: TO OTHER GLUCOSYLTRANSFERASES AND SOME TO A GLUCAN-				
CC	BINDING PROTEIN FROM S.MUTANS.				
CC	-1- SIMILARITY: Contains 16 cell wall binding repeats.				
CC	-----				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration				
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -				
CC	the European Bioinformatics Institute. There are no restrictions on its				
CC	use by non-profit institutions as long as its content is in no way				
CC	modified and this statement is not removed. Usage by and for commercial				
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/				
CC	or send an email to license@isb-sib.ch).				
CC	-----				
DR	EMBL; D90213; BA14241.1; -				Q9X616 macaca fasc
DR	InterPro; IPR002479; CW binding.				P11311 mycoplasma
DR	InterPro; IPR003318; Glyco hydro_70.				P21530 hog cholera
DR	Pfam; PF01473; CW binding_I: 13.				Q9S4T0 leghemoglobin
DR	Pfam; PF02324; Glyco hydro_70; 1.				P51136 dictyostella
FT	TRANSFERASE; Glycosyltransferase; Signal; Repeat; Dental caries.				P39668 saccharomyc
FT	SIGNAL				Q9Y7X6 schizosach
FT	CHAIN				001374 schizosoma
FT	DOMAIN				P51134 nicotiana t
FT					P35757 haemophilus
FT					P45991 haemophilus
FT					P75265 mycoplasma

```

FT DOMAIN 1093 1592 GLUCAN-BINDING (APPROXIMATE).
FT DOMAIN 1093 1592 6.5 X TANDEM REPEATS.
FT REPEAT 1093 1142 1.
FT REPEAT 1093 1142 2.
FT REPEAT 1222 1272 3.
FT REPEAT 1287 1337 4.
FT REPEAT 1402 1451 5.
FT REPEAT 1514 1563 6.
FT REPEAT 1577 1592 7 (INCOMPLETE).
SQ SEQUENCE 1592 AA; 176167 MW; BCOA6D079351ECF CRC64;

Query Match 100.0%; Score 119; DB 1; Length 1592;
Best Local Similarity 100.0%; Pred. No. 1,1e-10;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ANNHVSIVEAMSDNDTPYLHD 21
Db 477 ANNHVSIVEAMSDNDTPYLHD 497

RESULT 2
GTF1_STRMO STANDARD; PRT; 1597 AA.
ID GTF1_STRMO
AC P11001;
DT 01-JUN-1989 (Rel. 11, Created)
DT 01-JUN-1989 (Rel. 11, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Glucosyltransferase-1 precursor (EC 2.4.1.5) (GTF-I) (Dextranucrase)
DE (Sucrose 6-glucosyltransferase).
GN GTFI.
OS Streptococcus downei (Streptococcus sobrinus).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
NCBI_TaxID=1317;
[1]
SEQUENCE FROM N.A.
STRAIN=MFE28;
MEDLINE=87308014; PubMed=3040686;
Ferrecci J.J., Gilpin M.L., Russell R.R.B.;
"Nucleotide sequence of a glucosyltransferase gene from Streptococcus
sobrinus MFE28."
J. Bacteriol. 169:4271-4278(1987).
-1- FUNCTION: PRODUCTION OF EXTRACELLULAR GLUCANS, THAT ARE THOUGHT
TO PLAY A KEY ROLE IN THE DEVELOPMENT OF THE DENTAL PLAQUE BECAUSE
OF THEIR ABILITY TO ADHERE TO SMOOTH SURFACES AND MEDIANE THE
AGGREGATION OF BACTERIAL CELLS AND FOOD DEBRIS.
-1- CATALYTIC ACTIVITY: Sucrose + {(1,6)-alpha-D-glucosyl} (N) = D-
fructose + {(1,6)-alpha-D-glucosyl} (N+1).
-1- SUBCELLULAR LOCATION: Secreted.
-1- MISCELLANEOUS: GTF-I SYNTHESIZES WATER-INSOLUBLE GLUCANS (ALPHA
1,3-LINKED GLUCOSE AND SOME 1,6 LINKAGES), GTF-S SYNTHESIZES
WATER-SOLUBLE GLUCANS (ALPHA 1,6-GLUCOSE). GTF-SI SYNTHESIZES BOTH
FORMS OF GLUCANS.
-1- SIMILARITY: TO OTHER GLUCOSYLTRANSFERASES AND SOME TO A GLUCAN-
BINDING PROTEIN FROM S. MUTANS.
-1- SIMILARITY: Contains 19 cell wall binding repeats.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
CC EMBL: M17391; AAC63063.1; -
CC InterPro: IPR002479; CM binding.
CC InterPro: IPR003318; Glyco_hydro_70.
CC Pfam: PF01473; CM binding_1_16.
CC Pfam: PF02324; Glyco_hydro_70; 1.
CC Transferrase; Glycosyltransferase; Signal; Repeat; Dental caries.
CC SIGNAL 1 38 POTENTIAL.

```

```

FT CHAIN 39 1597 GLUCOSYLTRANSFERASE-I.
FT DOMAIN 39 1050 CATALYTIC (APPROXIMATE).
FT DOMAIN 1099 1597 GLUCAN-BINDING (APPROXIMATE).
FT REPEAT 1099 1597 1.25 A, 2 B, AND 5 AC REPEATS.
FT REPEAT 1099 1597 A REPEAT.
FT REPEAT 1163 1213 AC REPEAT.
FT REPEAT 1227 1277 AC REPEAT.
FT REPEAT 1292 1342 AC REPEAT.
FT REPEAT 1352 1399 B REPEAT.
FT REPEAT 1406 1455 AC REPEAT.
FT REPEAT 1465 1512 B REPEAT.
FT REPEAT 1519 1568 AC REPEAT.
FT REPEAT 1582 1597 A REPEAT (INCOMPLETE).
SQ SEQUENCE 1597 AA; 177080 MW; B9E86A200868798E CRC64;

Query Match 100.0%; Score 119; DB 1; Length 1597;
Best Local Similarity 100.0%; Pred. No. 1,1e-10;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ANNHVSIVEAMSDNDTPYLHD 21
Db 483 ANNHVSIVEAMSDNDTPYLHD 503

RESULT 3
GTFB_STRMU STANDARD; PRT; 1476 AA.
ID GTFB_STRMU
AC P08987; O69381; O69384; O69387; O69390; O69396;
DT 01-NOV-1988 (Rel. 09, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Glucosyltransferase-1 precursor (EC 2.4.1.5) (GTF-I) (Dextranucrase)
DE (Sucrose 6-glucosyltransferase).
GN GTFB OR SMU 1004.
OS Streptococcus mutans.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
NCBI_TaxID=1309;
[1]
SEQUENCE FROM N.A.
STRAIN=GS-5;
MEDLINE=87308013; PubMed=3040685;
Shirota T., Ueda S., Kuramitsu H.K.;
"Sequence analysis of the gtfB gene from Streptococcus mutans."
J. Bacteriol. 169:4263-4270(1987).
[2]
SEQUENCE FROM N.A.
STRAIN=MT4239 / Serotype C, MT4245 / Serotype E, MT4251 / Serotype F,
RC MT4467 / Serotype E, and MT8148 / Serotype C;
RX MEDLINE=98231643; PubMed=9570124;
Fujiwara T., Terao Y., Hoshino T., Kawabata S., Oshima T., Sobue S.,
Kimura S., Hamada S.;
"Molecular analyses of glucosyltransferase genes among strains of
Streptococcus mutans."
J. FEMS Microbiol. Lett. 161:331-336(1998).
[3]
SEQUENCE FROM N.A.
STRAIN=UA159 / ATCC 700610 / Serotype C;
RX MEDLINE=22295063; PubMed=12397186;
Aydin D., McShan W.M., McLaughlin R.E., Savic G., Chang J.,
Carzon W.B., Primeaux C., Tian R., Kenton S., Jia H., Lin S., Qian Y.,
Li S., Zhu H., Najjar F., Lai H., White J., Roe B.A., Ferretti J.J.;
"Genome sequence of Streptococcus mutans UA159, a cariogenic dental
pathogen."
Proc. Natl. Acad. Sci. U.S.A. 99:14434-14439(2002).
-1- FUNCTION: PRODUCTION OF EXTRACELLULAR GLUCANS, THAT ARE THOUGHT
TO PLAY A KEY ROLE IN THE DEVELOPMENT OF THE DENTAL PLAQUE BECAUSE
OF THEIR ABILITY TO ADHERE TO SMOOTH SURFACES AND MEDIANE THE
AGGREGATION OF BACTERIAL CELLS AND FOOD DEBRIS.
-1- CATALYTIC ACTIVITY: Sucrose + {(1,6)-alpha-D-glucosyl} (N) = D-
fructose + {(1,6)-alpha-D-glucosyl} (N+1).
-1- SUBCELLULAR LOCATION: Secreted.
-1- DISEASE: DENTAL CARIES.

```


CC -1- CATALYTIC ACTIVITY: Sucrose + {(1,6)-alpha-D-glucosyl}(N) = D-fructose + {(1,6)-alpha-D-glucosyl}(N+1).

CC -1- SUBCELLULAR LOCATION: Secreted.

CC -1- DISEASE: DENTAL CARIES.

CC -1- MISCELLANEOUS: GTF-1 SYNTHESIZES WATER-INSOLUBLE GLUCANS (ALPHA 1,3-LINKED GLUCOSE AND SOME 1,6 LINKAGES), GTF-S SYNTHESIZES WATER-SOLUBLE GLUCANS (ALPHA 1,6-GLUCOSE). GTF-SI SYNTHESIZES BOTH FORMS OF GLUCANS.

CC -1- SIMILARITY: TO OTHER GLUCOSYLTRANSFERASES AND SOME TO A GLUCAN-BINDING PROTEIN FROM S. MUTANS.

CC -1- SIMILARITY: Contains 6 cell wall binding repeats.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----

DR EMBL: M29296; AAA26895.1; -

DR EMBL: D88653; BAA26103.1; -

DR EMBL: D88656; BAA26107.1; -

DR EMBL: D88659; BAA26111.1; -

DR EMBL: D88662; BAA26115.1; -

DR EMBL: D89979; BAA26121.1; -

DR EMBL: AF014932; AAN58619.1; -

DR InterPro: IPR002479; CW_binding.

DR InterPro: IPR003318; Glyco_hydro_70.

DR Pfam: PF02324; Glyco_hydro_70; 1.

KW Transferase; Glycosyltransferase; Signal; Repeat; Dental caries; Complete proteome.

KM

FT SIGNAL 1 ?

FT CHAIN 1 1462

FT DOMAIN 1332 1423

FT REPEAT 1332 1295

FT REPEAT 1296 1359

FT REPEAT 1360 1423

FT REPEAT 10 10

FT VARIANT 19 19

FT VARIANT 58 58

FT VARIANT 68 68

FT VARIANT 81 81

FT VARIANT 113 113

FT VARIANT 122 122

FT VARIANT 132 132

FT VARIANT 135 135

FT VARIANT 137 137

FT VARIANT 202 202

FT VARIANT 255 255

FT VARIANT 275 275

FT VARIANT 288 288

FT VARIANT 301 301

FT VARIANT 313 313

FT VARIANT 317 317

FT VARIANT 328 328

FT VARIANT 350 350

FT VARIANT 628 633

FT VARIANT 688 688

FT VARIANT 726 730

FT VARIANT 762 762

FT VARIANT 964 964

FT VARIANT 1019 1019

Q -> H (IN STRAINS GS-5, MT4239, MT4245, MT4251, MT4467 AND MT8148).

I -> V (IN STRAINS GS-5, MT4239, MT4245, MT4251, MT4467 AND MT8148).

K -> E (IN STRAIN MT4467).

K -> S (IN STRAINS MT4239 AND MT4245).

A -> T (IN STRAINS MT4251 AND MT8148).

T -> I (IN STRAINS MT4239 AND MT4245).

A -> V (IN STRAINS MT4239, MT4245 AND MT8148).

A -> S (IN STRAINS GS-5 AND MT4467).

A -> V (IN STRAIN MT4245).

A -> T (IN STRAINS GS-5, MT4239, MT4245, MT4251, MT4467 AND MT8148).

V -> L (IN STRAIN MT4239).

V -> L (IN STRAIN MT4239).

D -> N (IN STRAIN MT8148).

E -> D (IN STRAINS MT4239, MT4245 AND MT4251).

D -> N (IN STRAINS MT4239, MT4245 AND MT4251).

Q -> H (IN STRAIN MT4245).

D -> N (IN STRAINS MT4239 AND MT4251).

E -> K (IN STRAIN MT4239).

V -> F (IN STRAIN MT4239).

F -> L (IN STRAINS MT4239, MT4251 AND MT4467).

KKKYYO -> EKEYTL (IN STRAIN MT4251).

A -> S (IN STRAIN MT4239).

IDOGSA -> ADKGNDS (IN STRAIN MT4251).

IDOGS -> ADKGN (IN STRAINS MT4239 AND MT4245).

T -> A (IN STRAINS GS-5, MT4239, MT4245, MT4251, MT4467 AND MT8148).

D -> Y (IN STRAIN MT4251).

E -> K (IN STRAINS MT4245 AND MT4251).

FT VARIANT 1059 1060

FT VARIANT 1060 1060

FT VARIANT 1080 1080

FT VARIANT 1142 1142

FT VARIANT 1198 1198

FT VARIANT 1220 1220

FT VARIANT 1280 1280

FT VARIANT 1282 1282

FT VARIANT 1290 1290

FT VARIANT 1311 1311

FT VARIANT 1403 1403

FT VARIANT 1425 1425

FT VARIANT 1449 1449

FT CONFLICT 1428 1462

FT SQ SEQUENCE 1462 AA; 163387 MW; CE4A279C4D708645 CRC64; R (IN REF. 1).

Query Match 62.2%; Score 74; DB 1; Length 1462;

Best Local Similarity 66.7%; Pred. No. 0.0015; Mismatches 5; Indels 0; Gaps 0;

Matches 14; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

OY 1 ANNHVSIVEAWSNDPTVYLHD 21

Db 495 AINHVSILEAWSNDPDQYMKD 515

RESULT 6

GTF5_STRDO STANDARD; PRT; 1365 AA.

AC P29336;

DT 01-DEC-1992 (Rel. 24, Created)

DT 01-DEC-1992 (Rel. 24, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Glucosyltransferase-5 precursor (EC 2.4.1.5). (GTF-S) (Dextranucrase)

DE (sucrose 6-glucosyltransferase).

GN GTF5.

OS Streptococcus downei (Streptococcus sobrinus).

OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;

OC Streptococcus.

OX NCBI_TaxID=13117;

PN [1]

PP SEQUENCE FROM N.A.

RC STRAIN=MFE28;

RX MEDLINE=90316665; PubMed=2142479;

RA Gilmore K.S., Russell R.R., Ferretti U.J.;

RT "Analysis of the Streptococcus downei gtf5 gene, which specifies a RT glucosyltransferase that synthesizes soluble glucans."

RL Infect. Immun. 58:2452-2458 (1990).

CC -1- FUNCTION: PRODUCTION OF EXTRACELLULAR GLUCANS, THAT ARE THOUGHT TO PLAY A KEY ROLE IN THE DEVELOPMENT OF THE DENTAL PLAQUE BECAUSE OF THEIR ABILITY TO ADHERE TO SMOOTH SURFACES AND MEDIANE THE AGGREGATION OF BACTERIAL CELLS AND FOOD DEBRIS.

CC -1- CATALYTIC ACTIVITY: Sucrose + {(1,6)-alpha-D-glucosyl}(N) = D-fructose + {(1,6)-alpha-D-glucosyl}(N+1).

CC -1- ENZYME REGULATION: GLUCAN SYNTHESIS BY GTF-S IS INDEPENDENT OF PRIMER GLUCAN UNLIKE GTF-1.

CC -1- DISEASE: DENTAL CARIES.

CC -1- MISCELLANEOUS: GTF-S SYNTHESIZES WATER-SOLUBLE GLUCANS (ALPHA 1,6-GLUCOSE).

CC -1- SIMILARITY: TO OTHER GLUCOSYLTRANSFERASES AND SOME TO A GLUCAN-BINDING PROTEIN FROM S. MUTANS.

CC -1- SIMILARITY: Contains 10 cell wall binding repeats.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----

DR EMBL: M30943; AAA26898.1; -

DR InterPro: IPR002479; CW_binding.

DR InterPro; IPR003318; Glyco_hydro_70.
DR Pfam; PF01473; CM_binding_1; 8.
DR Pfam; PF03324; Glycosyltransferase; Signal; Repeat; Dental caries.
KW Transferase; Glycosyltransferase; Signal; Repeat; Dental caries.
FT SIGNAL 1 36 OR 37 (POTENTIAL).
FT CHAIN 37 1365 GLYCOSYLTRANSFERASE-S.
FT REPEAT 157 177 CELL WALL BINDING 1.
FT REPEAT 178 197 CELL WALL BINDING 2.
FT DOMAIN 198 1061 CATALYTIC (APPROXIMATE).
FT REPEAT 1062 1082 CELL WALL BINDING 3.
FT REPEAT 1083 1102 CELL WALL BINDING 4.
FT REPEAT 1150 1169 CELL WALL BINDING 5.
FT REPEAT 1170 1190 CELL WALL BINDING 6.
FT REPEAT 1225 1243 CELL WALL BINDING 7.
FT REPEAT 1289 1308 CELL WALL BINDING 8.
FT REPEAT 1309 1328 CELL WALL BINDING 9.
FT REPEAT 1331 1352 CELL WALL BINDING 10.
SQ SEQUENCE 1365 AA; 151590 MW; 167296B5A2E8C476 CRC64;
Query Match 56.3%; Score 67; DB 1; Length 1365;
Best Local Similarity 57.1%; Pred. No. 0.018; Indels 0; Gaps 0;
Matches 12; Conservative 4; Mismatches 5;
QY 1 ANNHVSIVEAMSDNDTPYLH 21
DB 467 AIDHLSILEAMSGNDNDYKQ 487
ID UL16_EBV STANDARD; PRT; 336 AA.
AC P03221;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE BGLF2 protein.
GN BGLF2.
OS Epstein-Barr virus (strain B95-8) (Human herpesvirus 4).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Gammaherpesvirinae; Lymphocryptovirus.
OX NCBI_TaxID=10377;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=84270667; PubMed=6087149;
RA Baer R., Barker A.T., Biggin M.D., Deininger P.L., Farrell P.J.,
Gibson T.J., Hatfull G., Hudson G.S., Satchwell S.C., Seguin C.,
Tufnell P.S., Barrett B.G.;
RA "DNA sequence and expression of the B95-8 Epstein-Barr virus genome.";
RT Nature 310:207-211(1984).
RL Nature 310:207-211(1984).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=92113548; PubMed=1662696;
RA Chen M.R., Hsu T.Y., Lin S.W., Chen J.Y., Yang C.S.;
RT "Cloning and characterization of cDNA clones corresponding to
transcripts from the BamHI G region of the Epstein-Barr virus genome
and expression of BGLF2.";
RL J. Gen. Virol. 72:3047-3055(1991).
CC -1- SIMILARITY: BELONGS TO FAMILY THAT GROUPS TOGETHER HSV-1 UL16,
HSV-6 ORF1L1, EBV-1 46, HCMV UL94, EBV BDFP2, AND VZV 44.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC
CC EMBL; M60514; AAA45871.1; -
DR EMBL; V01555; CAA24831.1; -
DR EMBL; S77132; AAB21113.1; -
DR PIR; C43044; Q0BE40.
DR InterPro; IPR004286; UL16_UL94.

DR Pfam; PF03044; UL16_UL94; 1.
KW Late protein.
SQ SEQUENCE 336 AA; 36888 MW; 840937A416D5584C CRC64;
Query Match 41.2%; Score 49; DB 1; Length 336;
Best Local Similarity 41.2%; Pred. No. 2.5;
Matches 7; Conservative 6; Mismatches 4; Indels 0; Gaps 0;
QY 1 ANNHVSIVEAMSDNDTP 17
DB 216 AGAHVNIIRGWTEDDSP 232
ID G3PC_EIME STANDARD; PRT; 330 AA.
AC Q01558;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Glyceraldehyde 3-phosphate dehydrogenase, cytosolic (EC 1.2.1.12)
DE (GAPDH).
GN GAPC.
OS Leishmania mexicana.
OC Eukaryota; Euklenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.
OX NCBI_TaxID=55655;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=SP. Mexicana;
RX MEDLINE=93063042; PubMed=1435864;
RA Hannaert V., Blaauw M., Kohl L., Allert S., Operdoes F.R.,
Michels P.A.M.;
RA "Molecular analysis of the cytosolic and glycosomal glyceraldehyde-3-
phosphate dehydrogenase in the Leishmania mexicana.";
Mol. Biochem. Parasitol. 55:115-126(1992).
CC -1- CATALYTIC ACTIVITY: D-glyceraldehyde 3-phosphate + phosphate +
NAD(+) = 3-phospho-D-glyceroyl phosphate + NADH.
CC -1- PATHWAY: Second phase of glycolysis; first step.
CC -1- SUBUNIT: Homotetramer.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: Belongs to the glyceraldehyde 3-phosphate
dehydrogenase family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC
CC EMBL; X65220; CAA46323.1; -
DR PIR; B48445; B48445.
DR HSP; P06977; IGAD.
DR InterPro; IPR000173; GAP_dhrogenase.
DR InterPro; IPR006424; GAPDH-I.
DR Pfam; PF00044; gpch; 1.
DR Pfam; PF02800; gpch C; 1.
DR PRINTS: PR00078; G3PDHGNASE.
DR TIGRPMs; TIGR01534; GAPDH-I; 1.
DR PROSITE; PS00071; GAPDH; 1.
KW Glycolysis; Oxidoreductase; NAD.
FT INIT MET 0
FT BINDING 148 148 GLYCERALDEHYDE 3-PHOSPHATE.
FT ACT SITE 175 175 ACTIVATES THIOL GROUP DURING CATALYSIS.
SQ SEQUENCE 330 AA; 35511 MW; ED4H6D8E8A207E1E CRC64;
Query Match 40.3%; Score 48; DB 1; Length 330;
Best Local Similarity 42.1%; Pred. No. 3.5;
Matches 8; Conservative 4; Mismatches 7; Indels 0; Gaps 0;
QY 2 NNHVSIVEAMSDNDTPYLH 20

DB 300 NDHFVKLVSWNDETGYSH 318

RESULT 9

YMS8_YEAST STANDARD; PRT; 313 AA.

AC 003695;

DT 01-NOV-1997 (Rel. 35, Created)

DT 15-SEP-2003 (Rel. 42, Last sequence update)

DE Hypothetical 35.0 kDa protein in PFK2-HPA1 intergenic region.

GN YMS206W OR YMS325.07.

OS Saccharomyces cerevisiae (Baker's yeast).

OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;

OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.

OX NCBI_TaxID=4932;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=SD88C / AB972;

RX PubMed=9169872;

RA Bowman S., Churcher C.M., Badcock K., Brown D., Chillingworth T.,

RA Connor R., Dedman K., Devlin K., Gentles S., Hamlin N., Hunt S.,

RA Jagers K., Iye G., Moule S., Odell C., Pearson D., Rajandream M.A.,

RA Rice P., Skelton J., Walsh S., Whitehead S., Barrett B.G.,

RT "The nucleotide sequence of Saccharomyces cerevisiae chromosome

RT XIII."

RL Nature 387:90-93(1997).

CC -1 SIMILARITY: SOME, TO YEAST YNR014W.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC EMBL; 248755; CAA8648.1; -.

DR PIR; S59448; S59448.

DR SGD; S0004819; YMR206W.

KW Hypothetical protein.

FT DOMAIN 3

FT DOMAIN 146 149 POLY-SER.

FT DOMAIN 246 252 POLY-GLN.

FT DOMAIN 246 252 POLY-SER.

FT SEQUENCE 313 AA; 35017 MW; 9D92BFD8382577F0 CRC64;

Query Match 38.7%; Score 46; DB 1; Length 313; Best Local Similarity 47.4%; Pred. No. 6.9; Mismatches 8; Indels 0; Gaps 0;

Matches 9; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

DB 2 NNHVSIVEASDNDTPYLH 20

DB 125 NNHGFARFSDSHPSLH 143

RESULT 10

G3P_MYCGE STANDARD; PRT; 337 AA.

AC P47E43;

DT 01-FEB-1996 (Rel. 33, Created)

DT 01-FEB-1996 (Rel. 33, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Glyceraldehyde 3-phosphate dehydrogenase (EC 1.2.1.12) (GAPDH).

GN GAPD OR GAP OR MG301.

OS Mycoplasma genitalium.

OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.

OX NCBI_TaxID=2097;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=ATCC 33530 / G-37;

RX MEDLINE=96026346; PubMed=7569993;

RA Frazer C.M., Gocayne J.D., White O., Adams M.D., Clayton R.A., Fleischmann R.D., Bult C.J., Kerlavage A.R., Sutton G., Kelley J.M.,

RA Fritchman J.L., Weidman J.F., Small K.V., Sandusky M., Fuhrmann J.L.,
RA Nguyen D.T., Uterback T.R., Sauder D.M., Phillips C.A., Merrick J.M.,
RA Tomb J.F., Dougherty B.A., Bolt K.F., Hu P.-C., Luster T.S.,
RA Peterson S.N., Smith H.O., Hutchison C.A. III, Venter J.C.;
RT "The minimal gene complement of Mycoplasma genitalium.";
RT Science 270:1397-403(1995).

RN [2]
RP SEQUENCE OF 1-81 AND 279-337 FROM N.A.

RC STRAIN=ATCC 33530 / G-37;

RX MEDLINE=94075230; PubMed=8253680;

RA Peterson S.N., Hu P.-C., Bolt K.F., Hutchison C.A. III;

RT "A survey of the Mycoplasma genitalium genome by using random

RT sequencing."

RL J. Bacteriol. 175:7918-7930(1993).

CC -1 CATALYTIC ACTIVITY: D-glyceraldehyde 3-phosphate + phosphate +

CC NAD(+) = 3-phospho-D-glyceroyl phosphate + NADH.

CC -1 PATHWAY: Second phase of glycolysis; first step.

CC -1 SUBUNIT: Homotetramer (By similarity).

CC -1 SUBCELLULAR LOCATION: Cytoplasmic.

CC -1 SIMILARITY: Belongs to the glyceraldehyde 3-phosphate

CC dehydrogenase family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC EMBL; U39710; AAC71523.1; -.

DR EMBL; U02213; AAD12507.1; -.

DR EMBL; U02178; AAD12463.1; -.

DR PIR; C64233; C64233.

DR HSSP; P17721; 1HDG.

DR TIGR; MG301; -.

DR InterPro; IPR000173; GAP_dhydrogenase.

DR InterPro; IPR006424; GAPDH-T.

DR Pfam; PF00044; gpdh; 1.

DR Pfam; PF02800; gpdh_C; 1.

DR PRINTS; PR00078; G3PDHDSGNASE.

DR TIGRPFAMs; TIGR01534; GAPDH-1; 1.

DR PROSITE; PS00071; GAPDH; 1.

KM Glycolysis; Oxidoreductase; NAD, Complete proteome.

FT BINDING 157 157 GLYCERALDEHYDE 3-PHOSPHATE

FT ACT_SITE 184 184 ACTIVATES THIOL GROUP DURING CATALYSIS

FT FT (BY SIMILARITY).

FT SEQUENCE 337 AA; 37097 MW; FA1EA1966687006B CRC64;

Query Match 37.8%; Score 45; DB 1; Length 337; Best Local Similarity 53.8%; Pred. No. 11; Mismatches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

DB 8 VEASDNDTPYLH 20

DB 313 VYAWYDNESSYVH 325

RESULT 11

SILF_MOUSE STANDARD; PRT; 569 AA.

AC Q920G3;

DT 28-FEB-2003 (Rel. 41, Created)

DT 28-FEB-2003 (Rel. 41, Last sequence update)

DT 15-SEP-2003 (Rel. 42, Last annotation update)

DE Sialic acid binding Ig-like lectin-F precursor (msiglec-F).

GN SIGLECF.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_TaxID=10090;

RN [1]

```

RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6; PubMed=11579105;
RX MEDLINE=21576254; Varki N.M., Varki A.;
RA Angata T., Hingorani R., Varki N.M., Varki A.;
RT "Cloning and characterization of a novel mouse Siglec-F:
RT differential evolution of the mouse and human (CD33) Siglec-3-related
RT gene clusters."
RL J. Biol. Chem. 276:45128-45136(2001).
CC -1- FUNCTION: Putative adhesion molecule that mediates sialic acid
CC dependent binding to cells. Preferentially binds to alpha2,3-
CC linked sialic acid. The sialic acid recognition site may be masked
CC by cis interactions with sialic acids on the same cell surface.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- TISSUE SPECIFICITY: Predominantly expressed by immature
CC monocyte/myeloid lineage cells in bone marrow. Also found at
CC lower levels in mature neutrophils and monocytes.
CC -1- DOMAIN: Contains an intracytoplasmic motif referred as
CC immunoreceptor tyrosine-based inhibitor motif (ITIM). This motif
CC is involved in downmodulation of cellular functions as the
CC termination of the immune response.
CC -1- SIMILARITY: Belongs to the immunoglobulin superfamily. SIGLEC
CC (SIALIC ACID BINDING IG-LIKE LECTIN) SUBFAMILY.
CC -1- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.
CC -1- SIMILARITY: Contains 2 immunoglobulin-like C2-type domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF293371; AAL11043.1; -
DR InterPro; IPR007110; IG_1like.
DR InterPro; IPR003598; IG_C2.
DR InterPro; IPR003006; IG_MHC.
DR Pfam; PF00047; Ig_2.
DR SMART; SM00408; IgC2; 1.
DR PROSITE; PS50835; IG_LIKE; 2.
KM Cell adhesion; lectin; Transmembrane; signal; glycoprotein;
KM Immunoglobulin domain; Repeat.
FT STGNL 1 16
FT CHAIN 17 568
FT DOMAIN 17 439
FT TRANSMEM 440 460
FT DOMAIN 461 569
FT DOMAIN 18 116
FT DOMAIN 139 224
FT DOMAIN 229 324
FT SITE 536 541
FT SITE 559 564
FT DISULFID 35 163
FT DISULFID 40 96
FT DISULFID 157 206
FT DISULFID 265 308
FT CARBOHYD 95 95
FT CARBOHYD 151 151
FT CARBOHYD 200 200
FT CARBOHYD 203 203
FT CARBOHYD 369 369
FT CARBOHYD 372 372
FT CARBOHYD 387 387
SQ SEQUENCE 569 AA; 61476 MW; 8093838090484FC1 CRC64;

```

```

Query Match 37.0%; Score 44; DB 1; Length 569;
Best Local Similarity 57.1%; Pred. No. 30;
Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

```

```

QY 1 ANNHVSIVEAMSDN 14
DB 397 SSWHRLSCAMSDN 410

```

```

RESULT 12
YQ3C_YEAST STANDARD; PRT; 1122 AA.
ID YQ3C_YEAST
AC P53280;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Hypothetical 128.8 kDa protein in PAS2-PRS5 intergenic region.
GN YGRI34W.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
CX NCB1_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RA van Dyck U., Skala J., de Weydiffe P., Purnelle B., Talia E.,
RA Nawrocki A., del Bino S., Goffeau A.;
RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Z72919; CA97147.1; -
DR PIR; S64443; S64443.
DR SGD; S0003366; CAF130.
DR GO; GO:0030014; C:CCR4-NOT complex; IPI.
DR GO; GO:0006357; P:regulation of transcription from Pol II pro. .; IPI.
KW Hypothetical protein.
SQ SEQUENCE 1122 AA; 128779 MW; BF3542816CC78490 CRC64;

```

```

Query Match 37.0%; Score 44; DB 1; Length 1122;
Best Local Similarity 46.7%; Pred. No. 68;
Matches 7; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

```

```

QY 1 ANNHVSIVEAMSDN 15
DB 110 SSWHRLSCAMSDN 124

```

```

RESULT 13
YQ04_CAEEL STANDARD; PRT; 1131 AA.
ID YQ04_CAEEL
AC P53015;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE Hypothetical 127.2 kDa protein F31E3.4 in chromosome III.
GN F31E3.4.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Pelodierinae; Caenorhabditis.
CX NCB1_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RA Favallo A.;
RA STRAIN=Bristol N2;
RA Favallo A.;
RL Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: TO YEAST PAN2 AND TO S.POMBE SPAC22G7.04.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----

```

```

DR EMBL; U40935; AAA81687.1; -.
DR PIR; T16217; T16217.
DR Wormpep; F31E3.4; CE01269.
DR InterPro; IPR006055; Exonuclease.
DR Pfam; PF00929; Exonuclease; 1.
DR SMART; SMO0479; EXOIII; 1.
DR HYPOTHETICAL protein.
SQ SEQUENCE 1131 AA; 127174 MW; DE5AABOC309E7DE CRC64;

Query Match 37.0%; Score 44; DB 1; Length 1131;
Best Local Similarity 36.8%; Pred. No. 69;
Matches 7; Conservative 5; Mismatches 7; Indels 0; Gaps 0;

Oy 3 NHVSIVEAMSDNPPVLD 21
Db 315 NHGSIWVFPADRDQPOVNE 333

RESULT 14
PGCV MOUSE STANDARD; PRT; 3358 AA.
AC 062059; 062058; Q9CUT0;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Versican core protein precursor (large fibroblast proteoglycan)
DE (Chondroitin sulfate proteoglycan core protein 2) (PG-M).
GN CSG2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniota; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_Taxid=10090;

[1]
RP SEQUENCE FROM N.A. (ISOFORM V0; V1 AND V2).
RC STRAIN=C57BL/6; and Swiss Webster; TISSUE=Brain;
RX MEDLINE=95112551; PubMed=7822336;
RA Ito K., Shimomura T., Zako M., Ujita M., Kimata K.;
RT "Multiple forms of mouse PG-M, a large chondroitin sulfate
proteoglycan generated by alternative splicing.";
RL J. Biol. Chem. 270:958-965(1995).

[2]
RP SEQUENCE FROM N.A. (ISOFORM V3).
RC STRAIN=C57BL/6;
RX MEDLINE=95181355; PubMed=7876137;
RA Zako M., Shimomura T., Ujita M., Ito K., Kimata K.;
RT "Expression of pg-M(V3), an alternatively spliced form of PG-M
without a chondroitin sulfate attachment in region in mouse and human
tissues.";
RL J. Biol. Chem. 270:3914-3918(1995).

[3]
RP SEQUENCE OF 1-1692 FROM N.A. (ISOFORM V1).
RC STRAIN=C57BL/6J; TISSUE=Skin;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arikawa T., Hara A., Fukunishi Y., Komono H., Adachi U., Fukuda S.,
RA Saito K., Iwano M., Nishi K., Kiyosawa H., Kondo S., Yamanka I.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fletschmann W., Gaasterland T., Gissi C., King B., Kochwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schmitt L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Blake J., Boffelli D., Fujino M., Aono H., Balarelli R., Barsh G.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli U., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,
RA Wyshak-Bois A., Yoshida K., Hasegawa Y., Kawaji H., Kohlski S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).

```

```

RN [4]
RP INTERACTION WITH FBLN1.
RX PubMed=10400671;
RA Aspegberg A., Adam S., Kostka G., Timpl R., Heinegaard D.;
RT "Fibulin-1 is a ligand for the C-type lectin domains of aggrecan and
versican.";
RL J. Biol. Chem. 274:20444-20449(1999).
CC -1- FUNCTION: May play a role in intercellular signaling and in
connecting cells with the extracellular matrix. May take part in
the regulation of cell motility, growth and differentiation. Binds
hyaluronic acid.
CC -1- SUBUNIT: Interacts with FBLN1.
CC -1- SUBCELLULAR LOCATION: Secreted; extracellular matrix.
CC -1- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=4;
CC Comment=Additional isoforms seem to exist;
CC Name=V0;
CC IsoId=Q62059-1; Sequence=Displayed;
CC Name=V1;
CC IsoId=Q62059-2; Sequence=VSP_003087, VSP_003088;
CC Name=V2;
CC IsoId=Q62059-3; Sequence=VSP_003089;
CC Name=V3;
CC IsoId=Q62059-4; Sequence=VSP_003087, VSP_003090;
CC -1- TISSUE SPECIFICITY: V2 is found only in brain.
CC -1- DEVELOPMENTAL STAGE: Disappears after the cartilage development.
CC -1- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.
CC -1- SIMILARITY: Contains 2 EGF-like domains.
CC -1- SIMILARITY: Contains 2 EGF-like domains.
CC -1- SIMILARITY: Contains 1 C-type lectin family domain.
CC -1- SIMILARITY: Contains 1 Sushi (SCR) domain.
CC -1- SIMILARITY: BELONGS TO THE AGGRECAN/VERSICAN PROTEOGLYCAN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@sb-sib.ch).
CC -----
DR EMBL; D16263; BAA03796.1; -.
DR EMBL; D28599; NOT_ANNOTATED_CDS.
DR EMBL; D32040; BAA06802.1; -.
DR EMBL; AK014525; BAB29411.1; -.
DR HSSP; P01132; IEPG.
DR MGD; MGI:102889; Capg2.
DR InterPro; IPR000152; Asx_hydroxyl.
DR InterPro; IPR000742; EGF_2.
DR InterPro; IPR001881; EGF_Ca.
DR InterPro; IPR006209; EGF_like.
DR InterPro; IPR007110; IG-like.
DR InterPro; IPR003599; IG.
DR InterPro; IPR003006; IG_MHC.
DR InterPro; IPR001304; Lectin_C.
DR InterPro; IPR000538; Link.
DR InterPro; IPR000436; Sushi_SCR_CCP.
DR Pfam; PF00008; EGF_1.
DR Pfam; PF00047; IG; 1.
DR Pfam; PF00059; lectin_C; 1.
DR Pfam; PF00084; sushi; 1.
DR Pfam; PF00193; Xlink; 2.
DR PRINTS; PR01265; LINKMODULE.
DR ProDom; PD000918; Link; 2.
DR SMART; SMO0032; CCP; 1.
DR SMART; SMO0034; CLECT; 1.
DR SMART; SMO0179; EGF_CA; 1.
DR SMART; SMO0409; IG; 1.
DR SMART; SMO0445; LINK; 2.
DR PROSITE; PS00010; ASX_HYDROXYL; 1.
DR PROSITE; PS00615; C_TYPE_LLECTIN_1; 1.
DR PROSITE; PS50041; C_TYPE_LLECTIN_2; 1.
DR PROSITE; PS00022; EGF_1; 2.

```

DR PROSITE; PS01186; EGF_2; 1.
 DR PROSITE; PS01187; EGF_CA; 1.
 DR PROSITE; PS50835; IG_Like; 1.
 DR PROSITE; PS01241; LINK; 2.
 KW Glycoprotein; Proteoglycan; Lectin; Extracellular matrix; Sushi;
 KW Signal; Repeat; EGF-like domain; Calcium; Immunoglobulin domain;
 KW Hyalurononic acid; Alternative splicing.
 FT SIGNAL 1 20
 FT CHAIN 1 3358
 FT DOMAIN 21 146
 FT DOMAIN 167 244
 FT DOMAIN 265 346
 FT DOMAIN 348 1308
 FT DOMAIN 1309 3052
 FT DOMAIN 3052 3088
 FT DOMAIN 3090 3126
 FT DOMAIN 3139 3253
 FT DOMAIN 3258 3316
 FT DOMAIN 44 130
 FT DISULFID 172 243
 FT DISULFID 196 217
 FT DISULFID 270 333
 FT DISULFID 294 315
 FT DISULFID 3056 3076
 FT DISULFID 3061 3076
 FT DISULFID 3078 3087
 FT DISULFID 3094 3105
 FT DISULFID 3099 3114
 FT DISULFID 3116 3125
 FT DISULFID 3132 3143
 FT DISULFID 3160 3252
 FT DISULFID 3228 3244
 FT DISULFID 3259 3302
 FT DISULFID 3268 3315
 FT CARBOHYD 57 57
 FT CARBOHYD 330 330
 FT CARBOHYD 351 351
 FT CARBOHYD 441 441
 FT CARBOHYD 807 807
 FT CARBOHYD 914 914
 FT CARBOHYD 951 951
 FT CARBOHYD 1305 1305
 FT CARBOHYD 1372 1372
 FT CARBOHYD 1679 1679
 FT CARBOHYD 2054 2054
 FT CARBOHYD 2244 2244
 FT CARBOHYD 2362 2362
 FT CARBOHYD 2627 2627
 FT CARBOHYD 3030 3030
 FT CARBOHYD 3332 3332
 FT CARBOHYD 3342 3342
 FT CARBOHYD 348 348
 FT VARSPLIC 349 1308
 FT VARSPLIC 1309 3052
 FT VARSPLIC 349 3052
 FT CONFLICT 126 126
 FT CONFLICT 348 348
 FT CONFLICT 1658 1658
 FT CONFLICT 1674 1680
 FT CONFLICT 3358 3358
 SO SEQUENCE 3358 AA; 366938 MM; 071B80026BC0762D CRC64;
 Query Match 37.0%; Score 44; DB 1; Length 3358;
 Best Local Similarity 50.0%; Pred. No. 2.6e+02;
 Matches 8; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

RESULT 15
 ID 66PD_ECOLI STANDARD; PRT; 491 AA.
 AC P22932; P78069; O60134; O60139;
 AD 01-AUG-1991 (Rel. 19, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Glucose-6-phosphate 1-dehydrogenase (EC 1.1.1.49) (66PD).
 GN zwf OR B1852 OR C2265.
 OS Escherichia coli, and
 OS Escherichia coli O6.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Escherichia.
 OC NOBI_TaxID=562, 217992;
 [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=K12;
 RX MEDLINE=91123224; PubMed=1704005;
 RA Rowley D.L., Wolf R.E. Jr.;
 RT "Molecular characterization of the Escherichia coli K-12 zwf gene
 encoding glucose 6-phosphate dehydrogenase.";
 RL J. Bacteriol. 173:968-977(1991).
 [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=K12 / MG1655;
 RX MEDLINE=97426617; PubMed=9278503;
 RA Blatner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
 RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
 RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
 RA Mau B., Shao Y.;
 RT "The complete genome sequence of Escherichia coli K-12.";
 RL Science 277:1453-1474(1997).
 [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=K12;
 RX MEDLINE=97251358; PubMed=9097040;
 RA Itoh T., Alba H., Baba T., Fujita K., Hayashi K., Inada T.,
 RA Isono K., Kasai H., Kimura S., Kitakawa M., Kitagawa M.,
 RA Makino K., Maki T., Mizobuchi K., Mori H., Mori T., Motomura K.,
 RA Nakade S., Nakamura Y., Nishimoto H., Nishio Y., Oshima T.,
 RA Saito N., Sankai G., Seki Y., Siyusundaram S., Tagami H.,
 RA Takeda J., Takemoto K., Mada C., Yamamoto Y., Horuchi T.;
 RA "A 460-kb DNA sequence of the Escherichia coli K-12 genome
 corresponding to the 40.1-50.0 min region on the linkage map.";
 RL DNA Res. 3:379-392(1996).
 [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN=O6:H1 / CPT073 / ATCC 700928;
 RX MEDLINE=22388234; PubMed=12471157;
 RA Welch R.A., Burland V., Plunkett G. III, Redford P., Roesch P.,
 RA Rasko D., Buckles E.L., Liou S.-R., Boutin A., Hackett J., Stroud D.,
 RA Mayhew G.F., Rose D.J., Zhou S., Schwartz D.C., Perna N.T.,
 RA Mobley H.L.T., Domeneberg M.S., Blatner F.R.;
 RT "Extensive mosaic structure revealed by the complete genome sequence
 of uropathogenic Escherichia coli.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:17020-17024(2002).
 [5]
 RP SEQUENCE OF 72-368 FROM N.A.
 RC STRAIN=various ECOR strains;
 RX MEDLINE=95064015; PubMed=7973728;
 RA Gutman D.S., Dykhuizen D.E.;
 RT "Clonal divergence in Escherichia coli as a result of recombination,
 not mutation.";
 RL Science 266:1380-1383(1994).
 [6]
 RP SEQUENCE OF 321-491 FROM N.A.
 RC STRAIN=K12;
 RX MEDLINE=93345818; PubMed=8344525;
 RA Carter A.T., Pearson B.M., Dickinson J.R., Lancashire W.B.;
 RT "Sequence of the Escherichia coli K-12 edd and eda genes of the
 Entner-Doudoroff pathway.";

```
RL Gene 130:155-156(1993).
CC -|- CATALYTIC ACTIVITY: D-glucose 6-phosphate + NADP(+) = D-glucono-
CC 1,5-lactone 6-phosphate + NADPH.
CC -|- PATHWAY: Pentose phosphate pathway, first step.
CC -|- SIMILARITY: BELONGS TO THE GLUCOSE-6-PHOSPHATE DEHYDROGENASE
CC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M55005; AAA2475.1; -.
CC EMBL; AE000279; AAC74922.1; -.
CC EMBL; D90828; BAA15660.1; -.
CC EMBL; AE016761; AAN80722.1; -.
CC EMBL; U13783; AAA57018.1; -.
CC EMBL; U13784; AAA57019.1; -.
CC EMBL; U13785; AAA57020.1; -.
CC EMBL; U13786; AAA57021.1; -.
CC EMBL; U13787; AAA57022.1; -.
CC EMBL; U13788; AAA57023.1; -.
CC EMBL; U13789; AAA57024.1; -.
CC EMBL; U13790; AAA57025.1; -.
CC EMBL; U13791; AAA57026.1; -.
CC EMBL; U13792; AAA57027.1; -.
CC EMBL; U13793; AAA57028.1; -.
CC EMBL; U13794; AAA57029.1; -.
CC EMBL; X63694; CAA45220.1; -.
CC PIR; D64947; D64947.
CC DR HSSP; P11411; IDPG.
CC DR SWISS-2DPAGE; P22992; COLI.
CC DR ECODBASE; F048_8; 6TH EDITION.
CC DR Ecogene; EG11221; zwf.
CC DR InterPro; IPR001282; G6PD.
CC DR Pfam; PF00479; G6PD_C; 1.
CC DR Pfam; PF02781; G6PD; 1.
CC DR PRINTS; PRO0079; G6PDHGRNAS.
CC DR ProDom; PD001129; G6PD; 1.
CC DR TIGRfam; TIGR00871; zwf. 1.
CC DR PROSITE; PS00069; G6P_DEHYDROGENASE; 1.
CC KW Oxidoreductase; NADP; Glucose metabolism; Complete proteome.
CC FT ACT SITE 181 BY SIMILARITY.
CC FT VARIANT 100 S -> N (IN STRAINS ECOR4 AND ECOL10).
CC FT CONFLICT 268 LKSLRIDRSNVREKTVGGYTAGFA -> PEVSSPHRPLQ
CC FT RTRKRRRTAYCVP (IN REF. 1).
CC FT RTRKRRRTAYCVP (IN REF. 1).
CC SQ SEQUENCE 491 AA; 55704 MW; 263F07D298EAFCD3 CRC64;

Query Match 36.6%; Score 43.5; DB 1; Length 491;
Best Local Similarity 69.2%; Pred. No. 30;
Matches 9; Conservative 1; Mismatches 2; Indels 1; Gaps 1;

QY 6 SIVEAMS-DNDTP 17
Db 451 SITEAMAMDNDA 463
```

Search completed: November 13, 2003, 09:45:32
Job time : 7.56398 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 13, 2003, 09:31:40 : Search time 31.6493 Seconds
(without alignments)
171.224 Million cell updates/sec

Title: US-09-290-049a-14
Perfect score: 119
Sequence: 1 ANNHVSIVEAWSNDPTPLHD 21

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

SPREMBL_23:*

- 1: sp_bacteria:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phage:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_virus:*
- 16: sp_bacterioph:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	119	100.0	1590	2	Q59983 streptococc
2	119	100.0	1590	2	Q55263 streptococc
3	76	63.9	1016	2	Q91CJ7 leucostoc
4	72	60.5	2835	2	Q8G9Q2 streptococc
5	68	57.1	1477	2	Q91A66 leucostoc
6	68	57.1	1508	2	Q9EZH5 leucostoc
7	68	57.1	1508	2	Q52224 leucostoc
8	68	57.1	1575	2	Q91CH3 streptococc
9	68	57.1	1577	2	Q54178 streptococc
10	67	56.3	1338	2	Q9WJ4 streptococc
11	65	54.6	1527	2	Q9ZAR4 leucostoc
12	65	54.6	1527	2	Q8KRE1 leucostoc
13	61	51.3	1449	2	Q8KRE1 streptococc
14	61	51.3	1449	2	Q55264 streptococc
15	61	51.3	1554	2	Q8KZ15 streptococc
16	60	50.4	1512	2	Q9WJ4 streptococc

17	57	47.9	1577	2	Q55265 streptococc
18	56	47.1	1518	2	Q00600 streptococc
19	55	46.2	1599	2	Q00599 streptococc
20	55	46.2	2057	2	Q9RE05 leucostoc
21	53	44.5	175	10	Q91MJ8 arabidopsi
22	53	44.5	195	10	Q81DM8 arabidopsi
23	51	42.9	1195	5	Q06459 thermus the
24	50	42.0	601	5	Q9VGR4 drosophila
25	50	42.0	630	5	Q81A43 drosophila
26	49	41.2	524	16	Q9KSI8 vibrio chol
27	48.5	40.8	282	10	Q8RWJ3 allium cepa
28	48	40.3	184	16	Q8KWE4 clostridium
29	48	40.3	380	3	Q8N073 magnaporthe
30	47	39.5	193	2	Q9S445 pseudomonas
31	46.5	39.1	212	2	Q9KX16 yersinia en
32	46.5	39.1	214	11	Q8C1P7 mus musculu
33	46.5	39.1	1604	4	Q8NFA0 homo sapien
34	46	38.7	169	5	Q8S889 encephalito
35	46	38.7	271	11	Q9QZB9 rattus norv
36	46	38.7	283	5	Q9VIP7 drosophila
37	46	38.7	283	5	Q8WSP4 drosophila
38	46	38.7	393	16	Q8Y2Q6 ralsonia s
39	46	38.7	490	16	Q06579 mycobacteri
40	46	38.7	601	16	Q05566 mycobacteri
41	46	38.7	810	5	Q8T3J2 drosophila
42	46	38.7	811	5	Q9VK54 drosophila
43	46	38.7	880	5	P91643 drosophila
44	45.5	38.2	418	17	Q97U15 sulfolobus
45	45	37.8	119	2	Q93RH1 streptococc

ALIGNMENTS

RESULT 1

059983 PRELIMINARY: PRT; 1590 AA.

ID Q59983

AC Q59983;

DT 01-NOV-1996 (TREMBLrel. 01, Created)

DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)

DT 01-OCT-2002 (TREMBLrel. 22, Last annotation update)

DE Glucosyltransferase-I precursor (EC 2.4.1.5).

GN GFPI.

OS Streptococcus sobrinus.

OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;

OC Streptococcus.

OX NCBI_Taxid=1310;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=OM2176;

RX MEDLINE=94146405; PubMed=8312602;

RA Sato S., Inoue M., Handa N., Aizawa Y., Isobe Y., Katayama T.;

RT "DNA sequence of the glucosyltransferase gene of serotype d Streptococcus sobrinus.";

RT Streptococcus sobrinus.";

RL DNA Seq. 4:19-27 (1993).

DR EMBL; D13858; BAA02976.1; -

DR InterPro; IPR002479; CW binding.

DR InterPro; IPR003318; Glyco_hydro_70.

DR Pfam; PF01473; Glyco_hydro_1; 16.

DR Pfam; PF02324; Glyco_hydro_70; 1.

RW Glycosyltransferase; Signal; Transferase.

FT SIGNAL 1 38

FT CHAIN 39 1590

FT SEQUENCE 1590 AA; 175955 MW; C3C83A57CF3C2B0E CRC64;

Query Match 100.0%; Score 119; DB 2; Length 1590;
Best Local Similarity 100.0%; Pred. No. 1.2e-09;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ANNHVSIVEAWSNDPTPLHD 21
|||||
DB 477 ANNHVSIVEAWSNDPTPLHD 497

RESULT 2
ID Q55263 PRELIMINARY; PRT: 1590 AA.
AC Q55263;
DT 01-NOV-1996 (TRENBLrel. 01, Created)
DT 01-JAN-1998 (TRENBLrel. 05, Last sequence update)
DT 01-OCT-2002 (TRENBLrel. 22, Last annotation update)
DE GTP-I
GN GLUCOSYLTRANSFERASE.
OS Streptococcus sobrinus.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1310;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 33478;
RA Sato S.;
RT "DNA and amino-acid sequences of water-insoluble-glucan synthetase produced from Streptococcus sobrinus ATCC 33478.";
RT Ann. Kagoshima Univ. Dental School 16:23-29(1996).
DR EMBL; D63570; BAA09792.1; -
DR InterPro; IPR002479; CW_binding.
DR InterPro; IPR003318; Glyco_hydro_70.
DR Pfam; PF01473; CW_binding_1; 15.
DR Pfam; PF02324; Glyco_hydro_70; 1.
SQ SEQUENCE 1590 AA; 176058 MW; 9DF7A3F2C6E4FD43 CRC64;

Query Match 100.0%; Score 119; DB 2; Length 1590;
Best Local Similarity 100.0%; Pred. No. 1.2e-05;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ANNHVSIVEAWSNDPTPYLHD 21
DB 477 ANNHVSIVEAWSNDPTPYLHD 497

RESULT 3
ID Q9LCU7 PRELIMINARY; PRT: 1016 AA.
AC Q9LCU7;
DT 01-OCT-2000 (TRENBLrel. 15, Created)
DT 01-OCT-2000 (TRENBLrel. 15, Last sequence update)
DT 01-OCT-2002 (TRENBLrel. 22, Last annotation update)
DE Dextranucrase.
GN DSRRT
OS Leuconostoc mesenteroides.
OC Bacteria; Firmicutes; Lactobacillales; Leuconostoc.
OX NCBI_TaxID=1245;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NRRL B-512F;
RX MEDLINE=20169623; PubMed=10705445;
RA Funane K., Mizuno K., Takahara H., Kobayashi M.;
RT "Gene encoding a dextranucrase-like protein in Leuconostoc mesenteroides NRRL B-512F.";
RL Biosci. Biotechnol. Biochem. 64:29-38(2000).
DR EMBL; AB020020; BAA90527.1; -
DR HSPF; P06278; 1VJ5
DR InterPro; IPR003318; Glyco_hydro_70.
DR Pfam; PF02324; Glyco_hydro_70; 1.
SQ SEQUENCE 1016 AA; 110344 MW; 8896EFD813CCCB47 CRC64;

Query Match 63.9%; Score 76; DB 2; Length 1016;
Best Local Similarity 61.9%; Pred. No. 0.0042;
Matches 13; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 ANNHVSIVEAWSNDPTPYLHD 21
DB 548 ANNHVSIVEAWSNDPTPYLHD 568

RESULT 4

Q869Q2
ID Q869Q2 PRELIMINARY; PRT: 2835 AA.
AC Q869Q2;
DT 01-MAR-2003 (TRENBLrel. 23, Created)
DT 01-MAR-2003 (TRENBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update)
DE Dextranucrase (EC 2.4.1.5) (Fragment).
GN DSRF.
OS Leuconostoc mesenteroides.
OC Bacteria; Firmicutes; Lactobacillales; Leuconostoc.
OX NCBI_TaxID=1245;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=2231661; PubMed=12270834;
RX Bozonnet S., Dolis-Lafaygue M., Fabre E., Pizut S., Renaud-Simeon M.,
RA Mondan P., Willemot R.M.;
RT "Molecular characterisation of DSR-E, an alpha-1,2 linkage
RT synthesising dextranucrase with two catalytic domains.";
RL J. Bacteriol. 184:5753-5761(2002).
DR EMBL; AJ430204; CAD22883.1; -
DR Transferase; Glycosyltransferase.
FT NON TER
SQ SEQUENCE 2835 AA; 313264 MW; D03262CDD735399D CRC64;

Query Match 60.5%; Score 72; DB 2; Length 2835;
Best Local Similarity 57.9%; Pred. No. 0.058;
Matches 11; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

QY 1 ANNHVSIVEAWSNDPTPYLHD 19
DB 557 ANNHVSIVEAWSNDPTPYLHD 575

RESULT 5
ID Q9L466 PRELIMINARY; PRT: 1477 AA.
AC Q9L466;
DT 01-OCT-2000 (TRENBLrel. 15, Created)
DT 01-OCT-2000 (TRENBLrel. 15, Last sequence update)
DT 01-OCT-2002 (TRENBLrel. 22, Last annotation update)
DE Dextranucrase (EC 2.4.1.5).
GN DSRF.
OS Leuconostoc mesenteroides.
OC Bacteria; Firmicutes; Lactobacillales; Leuconostoc.
OX NCBI_TaxID=1245;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NRRL B-1355;
RA Arguello-Morales M.A., Renaud-Simeon M., Pizut S., Sarcabal P.,
RT "Sequence analysis of the gene encoding alternansucrase, a sucrose
RT glucosyltransferase from Leuconostoc mesenteroides NRRL B-1355.";
RT Submitted (OCT-1999) to the EMBL/Genbank/DBSJ databases.
DR EMBL; AJ250172; CAB76565.1; -
DR InterPro; IPR002479; CW_binding.
DR InterPro; IPR003318; Glyco_hydro_70.
DR Pfam; PF01473; CW_binding_1; 14.
DR Pfam; PF02324; Glyco_hydro_70; 1.
DR Pfam; PF02324; Glyco_hydro_70; 1.
KW Glycosyltransferase; Transferase.
SQ SEQUENCE 1477 AA; 164887 MW; B6F5710DEDFC831 CRC64;

Query Match 57.1%; Score 68; DB 2; Length 1477;
Best Local Similarity 57.1%; Pred. No. 0.12;
Matches 12; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 1 ANNHVSIVEAWSNDPTPYLHD 21
DB 532 ANNHVSIVEAWSNDPTPYLHD 552

RESULT 6
Q9EZHS PRELIMINARY; PRT: 1508 AA.

```

AC Q9BZ45;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Dextranucrase DsrB742.
GN DSRB742.
OS Leuconostoc mesenteroides.
OC Bacteria; Firmicutes; Lactobacillales; Leuconostoc.
OX NCBI_TaxID=1245;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=B-742CB;
RA Kim H.-S., Kim D., Ryu H.-J., Rohyt J.F.;
RT "Leuconostoc mesenteroides B-742CB, a dextranucrase gene.";
RL Submitted (Aug-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF294469; AAG38021.1; -.
DR InterPro; IPR002479; CW binding.
DR InterPro; IPR003318; Glyco_hydro_70.
DR Pfam; PF01473; CW binding_1; 14.
DR Pfam; PF02324; Glyco_hydro_70; 1.
SQ SEQUENCE 1508 AA; 168542 MW; 1.E2FCFA0F87A84F3A CRC64;

Query Match 57.1%; Score 68; DB 2; Length 1508;
Best Local Similarity 57.1%; Pred. No. 0.12;
Matches 12; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 1 ANNHVSIVEANSNDTPYIHD 21
Db 563 ANOHLSTLEDSHNDPEYKD 583

RESULT 7
052224 PRELIMINARY; PRT; 1508 AA.
ID 052224;
AC 052224;
DT 01-JUN-1998 (TrEMBLrel. 06, Created)
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Glucosyltransferase (EC 2.4.1.5).
GN DSRB.
OS Leuconostoc mesenteroides.
OC Bacteria; Firmicutes; Lactobacillales; Leuconostoc.
OX NCBI_TaxID=1245;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NRL B-1299;
RA Monchois V., Renaud-Simeon M., Monsan P., Willemot R.M.;
RT "Cloning and sequencing of a gene coding for an extracellular
RT dextranucrase (DSRB) from Leuconostoc mesenteroides NRRL B-1299
RT synthesizing only a a(1-6) glucan.";
RL EMBL; AF030129; AAB95453.1; -.
DR InterPro; IPR002479; CW binding.
DR InterPro; IPR003318; Glyco_hydro_70.
DR Pfam; PF01473; CW binding_1; 14.
DR Pfam; PF02324; Glyco_hydro_70; 1.
DR Pfam; PF02324; Glyco_hydro_70; 1.
KW Glycosyltransferase; transferase.
SQ SEQUENCE 1508 AA; 168511 MW; E70CECB57A70D1F0 CRC64;

Query Match 57.1%; Score 68; DB 2; Length 1508;
Best Local Similarity 57.1%; Pred. No. 0.12;
Matches 12; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 1 ANNHVSIVEANSNDTPYIHD 21
Db 563 ANOHLSTLEDSHNDPEYKD 583

RESULT 8
09LCH3 PRELIMINARY; PRT; 1575 AA.
ID 09LCH3;
AC 09LCH3;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)

```

```

DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Glucosyltransferase.
GN GTRF.
OS Streptococcus oralis.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1303;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC10557;
RX MEDLINE=20231779; PubMed=10768934;
RA Fujiwara T., Hoshino T., Ooshima T., Sobue S., Hamada S.;
RT "Purification, characterization, and molecular analysis of the gene
RT encoding glucosyltransferase from Streptococcus oralis.";
RL Infect. Immun. 68:2475-2483(2000).
DR EMBL; AB025228; BA095201.1; -.
DR InterPro; IPR002479; CW binding.
DR InterPro; IPR003318; Glyco_hydro_70.
DR Pfam; PF01473; CW binding_1; 17.
DR Pfam; PF02324; Glyco_hydro_70; 1.
DR Pfam; PF02324; Glyco_hydro_70; 1.
KW Transferase.
SQ SEQUENCE 1575 AA; 176792 MW; 772A26EAD7C2E543 CRC64;

Query Match 57.1%; Score 68; DB 2; Length 1575;
Best Local Similarity 61.9%; Pred. No. 0.13;
Matches 13; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 1 ANNHVSIVEANSNDTPYIHD 21
Db 546 AIKHLSTLEDSHNDPEYKD 566

RESULT 9
054178 PRELIMINARY; PRT; 1577 AA.
ID 054178;
AC 054178;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Glucosyltransferase.
GN GTRF.
OS Streptococcus gordonii Challis.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=29390;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CHALLIS;
RX MEDLINE=96157084; PubMed=8586195;
RA Vickerman M.M., Sulavik M.C., Clewell D.B.;
RT "Molecular analysis of Streptococcus gordonii glucosyltransferase
RT phase variants.";
RL Dev. Biol. Stand. 85:309-314(1995).
RN [2]
RP SEQUENCE OF 1-96 FROM N.A.
RC STRAIN=CHALLIS;
RX MEDLINE=92276317; PubMed=1534326;
RA Sulavik M.C., Tardif G., Clewell D.B.;
RT "Identification of a gene, rgg, which regulates expression of
RT glucosyltransferase and influences the spp phenotype of Streptococcus
RT gordonii Challis.";
RL J. Bacteriol. 174:3577-3586(1992).
DR EMBL; U12643; AAC43483.1; -.
DR EMBL; M89776; AAZ6969.1; -.
DR InterPro; IPR002479; CW binding.
DR InterPro; IPR003318; Glyco_hydro_70.
DR Pfam; PF01473; CW binding_1; 18.
DR Pfam; PF02324; Glyco_hydro_70; 1.
DR Pfam; PF02324; Glyco_hydro_70; 1.
KW Transferase.
SQ SEQUENCE 1577 AA; 177805 MW; 5AE0328DC5E08D18 CRC64;

Query Match 57.1%; Score 68; DB 2; Length 1577;

```


Best Local Similarity 61.9%; Pred. No. 0.13;
Matches 13; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 1 ANNHVSIVEAWSNDNTPYLHD 21
Db 548 ALKHLSTLEAWSNDNPDYKND 568

RESULT 10

Q9WXA4 PRELIMINARY; PRT; 1338 AA.
AC Q9WXA4;
DT 01-NOV-1999 (TREMBlrel. 12, Created)
DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
DE GTF-S.
GN GTF-S.
OS Streptococcus criceti.
OC Bacteri; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1333;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HS-6;
RA Inoue M., Fukui K., Miyagi A.;
RT "Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases."
RL EMBL; AB026123; BAAV7236.1;
DR InterPro; IPR002479; CW binding.
DR InterPro; IPR003318; Glyco_hydro_70.
DR Pfam; PF01473; CW binding_1; 10.
DR Pfam; PF02324; Glyco_hydro_70; 1.
KW Plasmaid.
SQ SEQUENCE 1338 AA; 148558 MW; 0A90C8E10E15D99B CRC64;

Query Match 56.3%; Score 67; DB 2; Length 1338;
Best Local Similarity 57.1%; Pred. No. 0.15;
Matches 12; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 1 ANNHVSIVEAWSNDNTPYLHD 21
Db 437 AIDHLSTLEAWSNDNPDYKND 457

RESULT 11

Q9ZAR4 PRELIMINARY; PRT; 1527 AA.
ID Q9ZAR4;
AC Q9ZAR4;
DT 01-MAY-1999 (TREMBlrel. 10, Created)
DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)
DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
DE Dextranucrase.
GN DEX.
OS Leuconostoc mesenteroides.
OC Bacteri; Firmicutes; Lactobacillales; Leuconostoc.
OX NCBI_TaxID=1245;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NRRL B-512-F;
RA Bhattacharjee R., Singh D.K.S.;
RT "Cloning and Molecular Characterization of Dextranucrase Gene from
leuconostoc mesenteroides NRRL B-512F."
RL Submitted (DEC-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; U81374; AAD10952.1;
DR InterPro; IPR002479; CW binding.
DR InterPro; IPR003318; Glyco_hydro_70.
DR Pfam; PF01473; CW binding_1; 16.
DR Pfam; PF02324; Glyco_hydro_70; 1.
KW SEQUENCE 1527 AA; 169709 MW; 1DFAFA237C743398 CRC64;

Query Match 54.6%; Score 65; DB 2; Length 1527;
Best Local Similarity 57.1%; Pred. No. 0.36;

Matches 12; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 1 ANNHVSIVEAWSNDNTPYLHD 21
Db 581 ANQHLSTLEAWSNDNPDYKND 601

RESULT 12

Q8KRE1 PRELIMINARY; PRT; 1527 AA.
AC Q8KRE1;
DT 01-OCT-2002 (TREMBlrel. 22, Created)
DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Dextranucrase Dard (EC 2.4.1.5).
GN DSRD
OS Leuconostoc mesenteroides.
OC Bacteri; Firmicutes; Lactobacillales; Leuconostoc.
OX NCBI_TaxID=1245;
RN [1]
RP SEQUENCE FROM N.A.
RA Neubauer H., Bauche A., Mollet B.;
RT "Isolation and characterization of the dextranucrase Dard of
leuconostoc mesenteroides loc4."
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY017384; AAG61158.1;
DR InterPro; IPR002479; CW binding.
DR InterPro; IPR003318; Glyco_hydro_70.
DR Pfam; PF01473; CW binding_1; 6.
DR Pfam; PF02324; Glyco_hydro_70; 1.
KW Transferase; Glycosyltransferase; F9P0DE220BD89668 CRC64;
SQ SEQUENCE 1527 AA; 169835 MW; F9P0DE220BD89668 CRC64;

Query Match 54.6%; Score 65; DB 2; Length 1527;
Best Local Similarity 57.1%; Pred. No. 0.36;
Matches 12; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 1 ANNHVSIVEAWSNDNTPYLHD 21
Db 581 ANQHLSTLEAWSNDNPDYKND 601

RESULT 13

O68542 PRELIMINARY; PRT; 1449 AA.
ID O68542;
AC O68542;
DT 01-AUG-1998 (TREMBlrel. 07, Created)
DT 01-AUG-1998 (TREMBlrel. 07, Last sequence update)
DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
DE Glucosyltransferase N (Fragment).
GN GTFN.
OS Streptococcus salivarius.
OC Bacteri; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1304;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=V1477;
RA Jaffe R.I.;
RT "Streptococcus salivarius V1477 gtfN."
RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF049609; AAC05156.1;
DR InterPro; IPR002479; CW binding.
DR InterPro; IPR003318; Glyco_hydro_70.
DR Pfam; PF01473; CW binding_1; 8.
DR Pfam; PF02324; Glyco_hydro_70; 1.
KW Transferase.
FT NON TER 1449 1449
SQ SEQUENCE 1449 AA; 159895 MW; 0700F6D748471BBB CRC64;

Query Match 51.3%; Score 61; DB 2; Length 1449;
Best Local Similarity 57.1%; Pred. No. 1.4;
Matches 12; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 1 ANNHVSTVEAMSDNDTPYLHD 21
 DB 536 AIRHLSITLWMSHNDAYNED 556

RESULT 14

ID Q55264 PRELIMINARY; PRT; 1449 AA.
 AC Q55264;
 DT 01-NOV-1996 (TREMBlrel. 01, Created)
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
 DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
 DE Glucosyltransferase precursor.
 GN GTF.
 OS Streptococcus salivarius.
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 OX NCBI_TaxID=1304;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=95122197; PubMed=7822030;
 RA Simpson C.L., Giffard P.M., Jacques N.A.;
 RT "Streptococcus salivarius ATCC 25975 possesses at least two genes
 coding for primer-independent glucosyltransferases.";
 RI Infect. Immun. 63:609-621(1995).
 DR EMBL; L35495; AAC41412.1; -;
 DR InterPro: IPR002479; CW_binding.
 DR InterPro: IPR003318; Glyco_hydro_70.
 DR Pfam; PF01473; CW_binding_1; 8.
 DR Pfam; PF02324; Glyco_hydro_70; 1.
 KW Signal; Transferase.
 KM SIGNAL 1 35 POTENTIAL.
 FT CHAIN 36 1449 GLUCOSYLTRANSFERASE.
 SQ SEQUENCE 1449 AA; 159984 MM; DD62F07306E86A46 CRC64;

Query Match 51.3%; Score 61; DB 2; Length 1449;
 Best Local Similarity 57.1%; Pred. No. 1.4;
 Matches 12; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 1 ANNHVSTVEAMSDNDTPYLHD 21
 DB 536 AIRHLSITLWMSHNDAYNED 556

RESULT 15

ID Q8KZL5 PRELIMINARY; PRT; 1554 AA.
 AC Q8KZL5;
 DT 01-OCT-2002 (TREMBlrel. 22, Created)
 DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)
 DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
 DE Glucosyltransferase.
 GN GTF.
 OS Streptococcus sobrinus.
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 OX NCBI_TaxID=1310;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21958684; PubMed=11960691;
 RA Hanada N., Fukushima K., Nomura Y., Sempuku H., Hayakawa M.,
 RA Mukasa H., Shiroza T., Abiko Y.;
 RT "Cloning and nucleotide sequence analysis of the Streptococcus
 sobrinus gtfV gene that produces a highly branched water-soluble
 glucan.";
 RT Blochim. Biophys. Acta 1570:75-79(2002).
 DR EMBL; AB089438; BAC07265.1; -;
 DR InterPro: IPR002479; CW_binding.
 DR InterPro: IPR003318; Glyco_hydro_70.
 DR Pfam; PF01473; CW_binding_1; 6.
 DR Pfam; PF02324; Glyco_hydro_70; 1.
 KW Transferase.

SQ SEQUENCE 1554 AA; 171676 MM; 6981BC1DAE24A73 CRC64;

Query Match 51.3%; Score 61; DB 2; Length 1554;
 Best Local Similarity 61.1%; Pred. No. 1.6;
 Matches 11; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 4 HVSTVEAMSDNDTPYLHD 21
 DB 488 HSTLWMSHNDAYNED 505

Search completed: November 13, 2003, 09:44:06
 Job time : 31.6493 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 13, 2003, 09:31:40 ; Search time 13.6351 Seconds
(without alignments)
65.165 Million cell updates/sec

Title: US-09-290-049A-14

Perfect score: 119
Sequence: 1 ANNHVSIVEAMSDNDTPYLMD 21

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:*

- 1: /cgn2_6/ptodata/2/1aa/5A_COMB.pep:*
- 2: /cgn2_6/ptodata/2/1aa/6A_COMB.pep:*
- 3: /cgn2_6/ptodata/2/1aa/6A_COMB.pep:*
- 4: /cgn2_6/ptodata/2/1aa/6B_COMB.pep:*
- 5: /cgn2_6/ptodata/2/1aa/PCTUS_COMB.pep:*
- 6: /cgn2_6/ptodata/2/1aa/backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	108	90.8	1475	3	US-09-007-999-2
2	108	90.8	1475	3	US-09-210-361-2
3	108	90.8	1475	4	US-09-740-274-2
4	99	83.2	1375	3	US-09-210-361-4
5	99	83.2	1375	4	US-09-740-274-4
6	74	62.2	545	4	US-09-604-957-4
7	74	62.2	1430	3	US-09-008-172-2
8	74	62.2	1430	3	US-09-210-361-6
9	74	62.2	1430	4	US-09-740-274-6
10	65	54.6	523	4	US-09-604-957-5
11	57	47.9	1577	2	US-08-793-824-2
12	55	46.2	584	4	US-09-604-957-6
13	55	46.2	2057	4	US-09-499-203-2
14	48	40.3	535	4	US-09-604-957-7
15	48	40.3	1278	4	US-09-604-957-3
16	47	39.5	749	4	US-09-562-737-96
17	43.5	36.6	1092	4	US-09-275-608-3
18	43	36.1	151	4	US-09-198-452A-815
19	43	36.1	234	4	US-09-134-001C-4514
20	43	36.1	647	4	US-09-134-001C-5458
21	42	35.3	200	2	US-08-836-442-4
22	42	35.3	287	4	US-09-252-991A-31548
23	42	35.3	327	4	US-09-107-532A-6181
24	42	35.3	334	6	5290690-11
25	42	35.3	404	4	US-09-252-991A-27960
26	42	35.3	454	4	US-09-252-991A-28000
27	42	35.3	455	3	US-09-362-473-6

28	42	35.3	529	2	US-08-836-442-3	Sequence 3, Appli
29	42	35.3	1627	1	US-07-665-792E-9	Sequence 9, Appli
30	42	35.3	3898	3	US-08-750-717-2	Sequence 2, Appli
31	41.5	34.9	69	4	US-09-252-991A-17383	Sequence 17383, A
32	41.5	34.9	130	4	US-09-328-352-4952	Sequence 4952, Ap
33	41	34.5	88	4	US-09-252-991A-21480	Sequence 21480, A
34	41	34.5	259	1	US-08-277-231A-3	Sequence 3, Appli
35	41	34.5	259	2	US-08-473-750-6	Sequence 6, Appli
36	41	34.5	259	2	US-08-477-326-6	Sequence 3, Appli
37	41	34.5	2465	2	US-08-596-291-3	Sequence 3, Appli
38	41	34.5	2465	3	US-09-100-804-3	Sequence 12, Appli
39	41	34.5	2466	3	US-09-080-855-12	Sequence 12, Appli
40	41	34.5	2466	4	US-09-566-076-12	Sequence 2, Appli
41	41	34.5	2466	5	PCT-US94-09943-2	Sequence 46, Appli
42	41	34.5	2485	3	US-09-290-640-46	Sequence 2, Appli
43	41	34.5	3898	2	US-08-876-991-2	Sequence 2, Appli
44	41	34.5	3898	2	US-09-059-853-2	Sequence 1, Appli
45	40.5	34.0	214	3	US-09-214-278-1	

ALIGNMENTS

```
RESULT 1
US-09-007-999-2
; Sequence 2, Application US/09007999
; Patent No. 6087559
; GENERAL INFORMATION:
; APPLICANT: Nicholas, Scott E.
; TITLE OF INVENTION: Substitutes for Modified Starch and
; FILE REFERENCE: 0356D
; CURRENT APPLICATION NUMBER: US/09/007,999
; CURRENT FILING DATE: 1998-01-16
; EARLIER APPLICATION NUMBER: 08/478,704
; EARLIER FILING DATE: 1995-06-07
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 1475
; TYPE: PRT
; ORGANISM: Streptococcus mutans
; US-09-007-999-2

Query Match      90.8%; Score 108; DB 3; Length 1475;
Best Local Similarity 85.7%; Pred. No. 1.3e-06;
Matches 18; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Cy      1 ANNHVSIVEAMSDNDTPYLMD 21
Db      481 ANNHVSIVEAMSDNDTPYLMD 501

RESULT 2
US-09-210-361-2
; Sequence 2, Application US/09210361
; Patent No. 6284479
; GENERAL INFORMATION:
; APPLICANT: Nicholas, Scott E.
; TITLE OF INVENTION: Substitutes for Modified Starches and
; FILE REFERENCE: 0357CR
; CURRENT APPLICATION NUMBER: US/09/210,361
; CURRENT FILING DATE: 1998-12-11
; EARLIER APPLICATION NUMBER: 09/007,999
; EARLIER FILING DATE: 1998-01-16
; EARLIER APPLICATION NUMBER: 08/478,704
; EARLIER FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 09/009,620
; EARLIER FILING DATE: 1998-01-20
; EARLIER APPLICATION NUMBER: 08/485,243
; EARLIER FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 09/008,172
```

```
; EARLIER FILING DATE: 1998-01-16
; EARLIER APPLICATION NUMBER: 08/482,711
; EARLIER FILING DATE: 1995-06-07
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 1475
; TYPE: PRT
; ORGANISM: Streptococcus mutans
; US-09-210-361-2

Query Match          90.8%; Score 108; DB 3; Length 1475;
Best Local Similarity 85.7%; Pred. No. 1.3e-08;
Matches 18; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY      1 ANNHVSIVEAMSDNDTPYLHD 21
        |||:|||||:|||||:|||||
DB      481 ANDHLSILEAMSDNDTPYLHD 501

RESULT 3
US-09-740-274-2
; Sequence 2, Application US/09740274
; Patent No. 6465203
; GENERAL INFORMATION:
; APPLICANT: Nichols, Scott E.
; TITLE OF INVENTION: Glucan-containing Compositions and Paper
; FILE REFERENCE: 0357CRD
; CURRENT APPLICATION NUMBER: US/09/740,274
; PRIOR FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: 09/210,361
; PRIOR FILING DATE: 1998-12-11
; PRIOR APPLICATION NUMBER: 09/007,999
; PRIOR FILING DATE: 1998-01-16
; PRIOR APPLICATION NUMBER: 08/478,704
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: 09/009,620
; PRIOR FILING DATE: 1998-01-20
; PRIOR APPLICATION NUMBER: 08/485,243
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: 09/008,172
; PRIOR FILING DATE: 1998-01-16
; PRIOR APPLICATION NUMBER: 08/482,711
; PRIOR FILING DATE: 1995-06-07
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 1475
; TYPE: PRT
; ORGANISM: Streptococcus mutans
; US-09-740-274-2

Query Match          90.8%; Score 108; DB 4; Length 1475;
Best Local Similarity 85.7%; Pred. No. 1.3e-08;
Matches 18; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY      1 ANNHVSIVEAMSDNDTPYLHD 21
        |||:|||||:|||||:|||||
DB      481 ANDHLSILEAMSDNDTPYLHD 501

RESULT 4
US-09-210-361-4
; Sequence 4, Application US/09210361
; Patent No. 6284479
; GENERAL INFORMATION:
; APPLICANT: Nichols, Scott E.
; TITLE OF INVENTION: Substitutes for Modified Starches and
; TITLE OF INVENTION: Latexes in Paper Manufacture
; FILE REFERENCE: 0357CR
; CURRENT APPLICATION NUMBER: US/09/210,361
; PRIOR FILING DATE: 1998-12-11
; PRIOR APPLICATION NUMBER: 09/007,999
; EARLIER APPLICATION NUMBER: 08/482,711
```

```
; EARLIER FILING DATE: 1998-01-16
; EARLIER APPLICATION NUMBER: 08/478,704
; EARLIER FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 09/009,620
; EARLIER FILING DATE: 1998-01-20
; EARLIER APPLICATION NUMBER: 08/485,243
; EARLIER FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 09/008,172
; EARLIER FILING DATE: 1998-01-16
; EARLIER APPLICATION NUMBER: 08/482,711
; EARLIER FILING DATE: 1995-06-07
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 1375
; TYPE: PRT
; ORGANISM: streptococcus mutans
; US-09-740-274-4

Query Match          83.2%; Score 99; DB 3; Length 1375;
Best Local Similarity 81.0%; Pred. No. 3.2e-07;
Matches 17; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY      1 ANNHVSIVEAMSDNDTPYLHD 21
        |||:|||||:|||||:|||||
DB      507 ANDHLSILEAMSYNDTPYLHD 527

RESULT 5
US-09-740-274-4
; Sequence 4, Application US/09740274
; Patent No. 6465203
; GENERAL INFORMATION:
; APPLICANT: Nichols, Scott E.
; TITLE OF INVENTION: Glucan-containing Compositions and Paper
; FILE REFERENCE: 0357CRD
; CURRENT APPLICATION NUMBER: US/09/740,274
; PRIOR FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: 09/210,361
; PRIOR FILING DATE: 1998-12-11
; PRIOR APPLICATION NUMBER: 09/007,999
; PRIOR FILING DATE: 1998-01-16
; PRIOR APPLICATION NUMBER: 08/478,704
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: 09/009,620
; PRIOR FILING DATE: 1998-01-20
; PRIOR APPLICATION NUMBER: 08/485,243
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: 09/008,172
; PRIOR FILING DATE: 1998-01-16
; PRIOR APPLICATION NUMBER: 08/482,711
; PRIOR FILING DATE: 1995-06-07
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 1375
; TYPE: PRT
; ORGANISM: streptococcus mutans
; US-09-740-274-4

Query Match          83.2%; Score 99; DB 4; Length 1375;
Best Local Similarity 81.0%; Pred. No. 3.2e-07;
Matches 17; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY      1 ANNHVSIVEAMSDNDTPYLHD 21
        |||:|||||:|||||:|||||
DB      507 ANDHLSILEAMSYNDTPYLHD 527

RESULT 6
US-09-604-957-4
; Sequence 4, Application US/09604957
; Patent No. 6466314
```

```

; GENERAL INFORMATION:
; APPLICANT: VAN GEEL-SCHUTTEN, GERRITDINA HENDRIKA
; APPLICANT: DIJKHUIZEN, LOBBERT
; APPLICANT: RAHMOUI, HAKIM
; APPLICANT: LEER, ROBERT-JAN
; TITLE OF INVENTION: NOVEL GLUCOSYLTRANSFERASE AND NOVEL GLUCAN
; FILE REFERENCE: EO 43388
; CURRENT APPLICATION NUMBER: US/09/604,957
; PRIOR APPLICATION NUMBER: 00201871.1
; PRIOR FILING DATE: 2000-05-25
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 545
; TYPE: PRT
; ORGANISM: Streptococcus mutans
; US-09-604-957-4

Query Match          62.2%; Score 74; DB 4; Length 545;
Best Local Similarity 66.7%; Pred. No. 0.0098;
Matches 14; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Cy 1 ANNHVSIYVAMSDNDPTYLHD 21
Db 75 AINHLSITLAWSDNDPQYNKD 95

RESULT 7
US-09-008-172-2
; Sequence 2, Application US/09008172
; Patent No. 6127602
; GENERAL INFORMATION:
; APPLICANT: Nichols, Scott E.
; TITLE OF INVENTION: Substitutes for Modified Starch and
; FILE REFERENCE: 0358D
; CURRENT APPLICATION NUMBER: US/09/008,172
; CURRENT FILING DATE: 1998-01-16
; EARLIER APPLICATION NUMBER: 08/482,711
; EARLIER FILING DATE: 1995-06-07
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 1430
; TYPE: PRT
; ORGANISM: Streptococcus mutans
; US-09-008-172-2

Query Match          62.2%; Score 74; DB 3; Length 1430;
Best Local Similarity 66.7%; Pred. No. 0.003;
Matches 14; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Cy 1 ANNHVSIYVAMSDNDPTYLHD 21
Db 495 AINHLSITLAWSDNDPQYNKD 515

RESULT 8
US-09-210-361-6
; Sequence 6, Application US/09210361
; Patent No. 6284479
; GENERAL INFORMATION:
; APPLICANT: Nichols, Scott E.
; TITLE OF INVENTION: Substitutes for Modified Starches and
; FILE REFERENCE: 0357CR
; CURRENT APPLICATION NUMBER: US/09/210,361
; CURRENT FILING DATE: 1998-12-11
; EARLIER APPLICATION NUMBER: 09/007,999
; EARLIER FILING DATE: 1998-01-16
; EARLIER APPLICATION NUMBER: 08/478,704
; EARLIER FILING DATE: 1995-06-07
```

```

; EARLIER APPLICATION NUMBER: 09/009,620
; EARLIER FILING DATE: 1998-01-20
; EARLIER APPLICATION NUMBER: 08/485,243
; EARLIER FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 09/008,172
; EARLIER FILING DATE: 1998-01-16
; EARLIER APPLICATION NUMBER: 08/482,711
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 6
; LENGTH: 1430
; TYPE: PRT
; ORGANISM: streptococcus mutans
; US-09-210-361-6

Query Match          62.2%; Score 74; DB 3; Length 1430;
Best Local Similarity 66.7%; Pred. No. 0.003;
Matches 14; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Cy 1 ANNHVSIYVAMSDNDPTYLHD 21
Db 495 AINHLSITLAWSDNDPQYNKD 515

RESULT 9
US-09-740-274-6
; Sequence 6, Application US/09740274
; Patent No. 6465203
; GENERAL INFORMATION:
; APPLICANT: Nichols, Scott E.
; TITLE OF INVENTION: Glucan-containing Compositions and Paper
; FILE REFERENCE: 0357CRD
; CURRENT APPLICATION NUMBER: US/09/740,274
; CURRENT FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: 09/210,361
; PRIOR FILING DATE: 1998-12-11
; PRIOR APPLICATION NUMBER: 09/007,999
; PRIOR FILING DATE: 1998-01-16
; PRIOR APPLICATION NUMBER: 08/478,704
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: 09/009,620
; PRIOR FILING DATE: 1998-01-20
; PRIOR APPLICATION NUMBER: 08/485,243
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: 09/008,172
; PRIOR FILING DATE: 1998-01-16
; PRIOR APPLICATION NUMBER: 08/482,711
; PRIOR FILING DATE: 1995-06-07
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 6
; LENGTH: 1430
; TYPE: PRT
; ORGANISM: streptococcus mutans
; US-09-740-274-6

Query Match          62.2%; Score 74; DB 4; Length 1430;
Best Local Similarity 66.7%; Pred. No. 0.003;
Matches 14; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Cy 1 ANNHVSIYVAMSDNDPTYLHD 21
Db 495 AINHLSITLAWSDNDPQYNKD 515

RESULT 10
US-09-604-957-5
; Sequence 5, Application US/09604957
; Patent No. 6486314
; GENERAL INFORMATION:
; APPLICANT: VAN GEEL-SCHUTTEN, GERRITDINA HENDRIKA
; APPLICANT: DIJKHUIZEN, LOBBERT
```

```

; APPLICANT: RAHAOUI, HAKIM
; APPLICANT: LEER, ROBERT-JAN
; TITLE OF INVENTION: NOVEL GLUCOSYLTRANSFERASE AND NOVEL GLUCAN
; FILE REFERENCE: BO 43388
; CURRENT APPLICATION NUMBER: US/09/604,957
; CURRENT FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: 00201871.1
; PRIOR FILING DATE: 2000-05-25
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 5
; LENGTH: 523
; TYPE: PRT
; ORGANISM: Leuconostoc mesenteroides
; US-09-604-957-5

Query Match          54.6%; Score 65; DB 4; Length 523;
Best Local Similarity 57.1%; Pred. No. 0.025;
Matches 12; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY      1 ANNHVSIVEAMSDNDTPYLHD 21
DB      75 ANQHLSTLEDMWNGKDPQYVN 95

RESULT 11
US-08-793-824-2
; Sequence 2, Application US/08793824
; Patent No. 5981838
; GENERAL INFORMATION:
; APPLICANT: Simpson, Christine Lynn
; APPLICANT: Giffard, Philip Morrison
; APPLICANT: Jacques, Nicholas Anthony
; TITLE OF INVENTION: Genetic Manipulation of plants to
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESSES:
; ADDRESSER: Griffith Hack & Co
; STREET: Level 8, 168 Walker Street
; CITY: No. 5981838th Sydney
; STATE: New South Wales
; COUNTRY: Australia
; ZIP: 2060
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/793,824
; FILING DATE:
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: AU PM7643
; FILING DATE: 24-AUG-1994
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 61 2 9957 5944
; TELEFAX: 61 2 957 6288
; TELEX: 26547
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1577 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: not relevant
; MOLECULE TYPE: protein
; ORIGINAL SOURCE:
; ORGANISM: Streptococcus salivarius
; US-08-793-824-2

Query Match          47.9%; Score 57; DB 2; Length 1577;
Best Local Similarity 61.1%; Pred. No. 1.6;
Matches 11; Conservative 2; Mismatches 5; Indels 0; Gaps 0;
```

```

QY      4 HVSIVEAMSDNDTPYLHD 21
DB      591 HSLTLEAMSYNDHQYKND 608

RESULT 12
US-09-604-957-6
; Sequence 6, Application US/09604957
; Patent No. 6486314
; GENERAL INFORMATION:
; APPLICANT: VAN GELT-SCHUTTEN, GERRITDINA HENDRIKA
; APPLICANT: DIKHOIZEN, LUBBERT
; APPLICANT: RAHAOUI, HAKIM
; APPLICANT: LEER, ROBERT-JAN
; TITLE OF INVENTION: NOVEL GLUCOSYLTRANSFERASE AND NOVEL GLUCAN
; FILE REFERENCE: BO 43388
; CURRENT APPLICATION NUMBER: US/09/604,957
; CURRENT FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: 00201871.1
; PRIOR FILING DATE: 2000-05-25
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 6
; LENGTH: 584
; TYPE: PRT
; ORGANISM: Leuconostoc mesenteroides
; US-09-604-957-6

Query Match          46.2%; Score 55; DB 4; Length 584;
Best Local Similarity 45.0%; Pred. No. 1.1;
Matches 9; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

QY      1 ANNHVSIVEAMSDNDTPYLH 20
DB      75 ANKHLSTLEDMWNGKDPQYVN 94

RESULT 13
US-09-499-203-2
; Sequence 2, Application US/09499203
; Patent No. 6570065
; GENERAL INFORMATION:
; APPLICANT: KOSSMANN, Jens
; APPLICANT: WELSH, Thomas
; APPLICANT: QUANZ, Martin
; APPLICANT: KNUTH, Karola
; TITLE OF INVENTION: Nucleic Acid Molecules Encoding Alternansucrase
; FILE REFERENCE: 147-196P
; CURRENT APPLICATION NUMBER: US/09/499,203
; CURRENT FILING DATE: 2000-02-08
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 2
; LENGTH: 2057
; TYPE: PRT
; ORGANISM: Leuconostoc mesenteroides
; US-09-499-203-2

Query Match          46.2%; Score 55; DB 4; Length 2057;
Best Local Similarity 45.0%; Pred. No. 4.5;
Matches 9; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

QY      1 ANNHVSIVEAMSDNDTPYLH 20
DB      665 ANKHLSTLEDMWNGKDPQYVN 684

RESULT 14
US-09-604-957-7
; Sequence 7, Application US/09604957
; Patent No. 6486314
; GENERAL INFORMATION:
```

```

1  APPLICANT: VAN GEEL-SCHUTTEN, GERTRIDINA HENDRIKA
2  APPLICANT: DIJKHUIZEN, LUBBERT
3  APPLICANT: RAHMOUI, HAKIM
4  APPLICANT: LEER, ROBERT-JAN
5  TITLE OF INVENTION: NOVEL GLUCOSYLTRANSFERASE AND NOVEL GLUCAN
6  FILE REFERENCE: BO 43388
7  CURRENT APPLICATION NUMBER: US/09/604,957
8  CURRENT FILING DATE: 2000-06-28
9  PRIOR APPLICATION NUMBER: 00201871.1
10 PRIOR FILING DATE: 2000-05-25
11 NUMBER OF SEQ ID NOS: 17
12 SOFTWARE: SeqIntn Ver. 2.1
13 SEQ ID NO 7
14 LENGTH: 535
15 TYPE: PR1
16 ORGANISM: Lactobacillus reuteri
17 US-09-604-957-7

```

Query Match	40.3%;	Score 48;	DB 4;	Length 535;
Best Local Similarity	35.0%;	Pred. No. 12;		
Matches	7;	Conservative	6;	Mismatches 7;
				Indels 0;
				Gaps 0;

```
QY      1 ANNHVSIVEAWSNDNTPYLH 20  
        :|:::||:|:|:  
Db     74 SNKHINILEDWNHADPEYFN 93
```

```

RESULT 15
US-09-604-957-3
; Sequence 3, Application US/09604957
; Patent No. 6486314
; GENERAL INFORMATION:
; APPLICANT: VAN GEEEL-SCHOUTEN, GERRITDINA HENDRIJA
; APPLICANT: DIJKHUIZEN, LOUBERT
; APPLICANT: RAHOUTI, HAKIM
; APPLICANT: LEER, ROBERT-JAN
; TITLE OF INVENTION: NOVEL GLUCOSYLTRANSFERASE AND NOVEL GLUCAN
; FILE REFERENCE: BO 43388
; CURRENT APPLICATION NUMBER: US/09/604,957
; CURRENT FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: 00201871.1
; PRIOR FILING DATE: 2000-05-25
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 1278
; TYPE: prt
; ORGANISM: Lactobacillus reuteri
US-09-604-957-3

```

Query Match	40.34	Score 48	DB 4	Length 1278
Best Local Similarity	35.04	Pred. No. 34		
Matches	7	Conservative	6	Mismatches 7
				Indels 0
				Gaps 0

```
QY      1 ANNHVSIVEAWSNDNDTPYLH 20  
       :|::||:|:|:|:|:|:|:  
Db     550 SNKHINILEDWNHADPEYFN 569
```

```
Search completed: November 13, 2003, 09:48:01
Job time : 13.6351 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 13, 2003, 09:45:40 / Search time 27.4692 Seconds
(without alignments)
139.565 Million cell updates/sec

Title: US-09-290-049A-14
Perfect score: 119
Sequence: 1 ANNHVSIVEAMSDNDTPYLHD 21

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 666188 seqs, 182559486 residues

Total number of hits satisfying chosen parameters: 666188

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database:

1: /cgn2_6/ptodata/2/pubppaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/2/pubppaa/PCR_NEW_PUB.pep:*
3: /cgn2_6/ptodata/2/pubppaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/2/pubppaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/2/pubppaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/2/pubppaa/PCRUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/2/pubppaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/2/pubppaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/2/pubppaa/US09_PUBCOMB.pep:*
10: /cgn2_6/ptodata/2/pubppaa/US09_PUBCOMB.pep:*
11: /cgn2_6/ptodata/2/pubppaa/US09_NEW_PUB.pep:*
12: /cgn2_6/ptodata/2/pubppaa/US10_PUBCOMB.pep:*
13: /cgn2_6/ptodata/2/pubppaa/US10_PUBCOMB.pep:*
14: /cgn2_6/ptodata/2/pubppaa/US10_PUBCOMB.pep:*
15: /cgn2_6/ptodata/2/pubppaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/2/pubppaa/US10C_NEW_PUB.pep:*
17: /cgn2_6/ptodata/2/pubppaa/US60_PUBCOMB.pep:*
18: /cgn2_6/ptodata/2/pubppaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	108	90.8	1475	9 US-09-740-274-2	Sequence 2, Appli
2	99	83.2	1375	9 US-09-740-274-4	Sequence 4, Appli
3	74	62.2	545	10 US-09-995-749A-10	Sequence 11, Appli
4	74	62.2	1430	9 US-09-740-274-6	Sequence 6, Appli
5	65	54.6	522	10 US-09-995-749A-11	Sequence 11, Appli
6	55	46.2	584	10 US-09-995-749A-12	Sequence 12, Appli
7	48	40.3	535	10 US-09-995-749A-13	Sequence 13, Appli
8	47	39.5	1781	10 US-09-995-749A-2	Sequence 2, Appli
9	48	40.3	749	15 US-10-211-962-96	Sequence 96, Appli
10	46.5	39.1	1604	9 US-09-888-615-73	Sequence 73, Appli
11	46	38.7	502	12 US-10-009-823A-1	Sequence 1, Appli
12	46	38.7	1494	15 US-10-195-144-17	Sequence 17, Appli
13	45	37.8	199	10 US-09-738-626-4571	Sequence 4571, Ap
14	45	37.8	665	15 US-10-156-761-14923	Sequence 14923, A
15	45	37.8	972	10 US-09-924-154-16	Sequence 16, Appli

16	44	37.0	15	12	US-10-295-693-90	Sequence 90, Appli
17	44	37.0	15	15	US-10-067-649-90	Sequence 90, Appli
18	43.5	36.6	202	10	US-09-280-197-11	Sequence 11, Appli
19	43.5	36.6	202	12	US-10-448-139-11	Sequence 11, Appli
20	43.5	36.6	491	9	US-09-815-242-10190	Sequence 10190, A
21	43.5	36.6	491	9	US-09-815-242-13756	Sequence 13756, A
22	43.5	36.6	1092	10	US-09-423-126-5	Sequence 5, Appli
23	43	36.1	27	15	US-10-093-908-26	Sequence 26, Appli
24	43	36.1	67	15	US-10-050-704-324	Sequence 324, App
25	43	36.1	87	15	US-10-050-704-149	Sequence 149, App
26	43	36.1	332	10	US-09-454-279-20	Sequence 20, Appli
27	43	36.1	437	11	US-09-374-046A-66	Sequence 66, Appli
28	43	36.1	539	11	US-10-230-026-34	Sequence 34, Appli
29	43	36.1	603	15	US-10-054-678-35	Sequence 2, Appli
30	43	36.1	603	15	US-10-092-908-35	Sequence 35, Appli
31	43	36.1	715	15	US-10-156-761-7562	Sequence 7562, Ap
32	42	35.3	349	12	US-10-220-381-22	Sequence 22, Appli
33	42	35.3	432	9	US-09-815-242-5096	Sequence 5096, Ap
34	42	35.3	455	9	US-09-742-954-6	Sequence 6, Appli
35	42	35.3	7349	12	US-10-314-657-46	Sequence 46, Appli
36	41.5	34.9	49	12	US-10-029-386-28981	Sequence 28981, A
37	41	34.5	120	14	US-10-001-835-146	Sequence 1404, A
38	41	34.5	302	9	US-09-815-242-14004	Sequence 9, Appli
39	41	34.5	358	11	US-09-095-478-7	Sequence 7, Appli
40	41	34.5	381	11	US-09-095-478-7	Sequence 14114, A
41	41	34.5	438	15	US-10-156-761-14114	Sequence 11729, A
42	41	34.5	749	15	US-10-156-761-11729	Sequence 56, Appli
43	41	34.5	1267	15	US-10-059-585-56	Sequence 12, Appli
44	41	34.5	2466	12	US-10-177-980-12	Sequence 46, Appli
45	41	34.5	2485	9	US-09-802-669-46	Sequence 46, Appli

ALIGNMENTS

RESULT 1
US-09-740-274-2
Sequence 2, Application US/09740274
Patent No. US20020031826A1

GENERAL INFORMATION:
APPLICANT: Nichols, Scott E.

TITLE OR INVENTION: Glucan-containing Compositions and Paper

FILE REFERENCE: 0357CRD
CURRENT APPLICATION NUMBER: US/09/740,274

PRIOR FILING DATE: 2000-12-19
PRIOR APPLICATION NUMBER: 09/210,361

PRIOR FILING DATE: 1998-12-11
PRIOR APPLICATION NUMBER: 09/007,999

PRIOR FILING DATE: 1998-01-16
PRIOR APPLICATION NUMBER: 08/478,704

PRIOR FILING DATE: 1995-06-07
PRIOR APPLICATION NUMBER: 09/009,620

PRIOR FILING DATE: 1998-01-20
PRIOR APPLICATION NUMBER: 08/485,243

PRIOR FILING DATE: 1995-06-07
PRIOR APPLICATION NUMBER: 09/008,172

PRIOR FILING DATE: 1998-01-16
PRIOR APPLICATION NUMBER: 08/482,711

PRIOR FILING DATE: 1995-06-07
NUMBER OF SEQ ID NOS: 6

SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 2

LENGTH: 1475
TYPE: PRT

ORGANISM: Streptococcus mutans
US-09-740-274-2

Query Match 90.8% Score 108; DB 9; Length 1475;
Best Local Similarity 85.7% Pred. No. 7.3e-09;
Matches 18; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

1 ANNHVSIVEAMSDNDTPYLHD 21
|||:|||||

Query Match	62.2%	Score 74	DB 10	Length 545
Best Local Similarity	66.7%	Pred. No. 0.0042		
Matches 14	Conservative	2	Mismatches 5	Indels 0
				Gaps 0

Query Match	54.6%;	Score 65;	DB 10;	Length 522,
Best Local Similarity	57.1%;	Pred. No. 0.097;		

Matches 12; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 1 ANNHVSIWEAMSDNDTPYLH 21
Db 75 ANKHSTLEDMNGNDPLVYTD 95

RESULT 6
US-09-995-749A-12
; Sequence 12, Application US/09995749A
; Patent No. US20020155568A1
; GENERAL INFORMATION:
; APPLICANT: VAN GEEL-SCHUTTEN, GERRITDINA HENDRIKA
; APPLICANT: DIJKHUIZEN, LUBBERT
; APPLICANT: RAHAOUI, HAKIM
; APPLICANT: LEER, ROBERT-JAN
; TITLE OF INVENTION: NOVEL GLUCOSYLTRANSFERASES
; FILE REFERENCE: BO43388-CIP
; CURRENT APPLICATION NUMBER: US/09/995,749A
; CURRENT FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: 09/604,957
; PRIOR FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: EPO 00201871.1
; PRIOR FILING DATE: 2000-05-25
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 12
; LENGTH: 584
; TYPE: PRT
; ORGANISM: Leuconostoc mesenteroides
US-09-995-749A-12

Query Match 46.2%; Score 55; DB 10; Length 584;
Best Local Similarity 45.0%; Pred. No. 3.8;
Matches 9; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

QY 1 ANNHVSIWEAMSDNDTPYLH 20
Db 75 ANKHSTLEDMNGNDPLVYTD 94

RESULT 7
US-09-995-749A-13
; Sequence 13, Application US/09995749A
; Patent No. US20020155568A1
; GENERAL INFORMATION:
; APPLICANT: VAN GEEL-SCHUTTEN, GERRITDINA HENDRIKA
; APPLICANT: DIJKHUIZEN, LUBBERT
; APPLICANT: RAHAOUI, HAKIM
; APPLICANT: LEER, ROBERT-JAN
; TITLE OF INVENTION: NOVEL GLUCOSYLTRANSFERASES
; FILE REFERENCE: BO43388-CIP
; CURRENT APPLICATION NUMBER: US/09/995,749A
; CURRENT FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: 09/604,957
; PRIOR FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: EPO 00201871.1
; PRIOR FILING DATE: 2000-05-25
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 13
; LENGTH: 535
; TYPE: PRT
; ORGANISM: Lactobacillus reuteri
US-09-995-749A-13

Query Match 40.3%; Score 48; DB 10; Length 535;
Best Local Similarity 35.0%; Pred. No. 4.1;
Matches 7; Conservative 6; Mismatches 7; Indels 0; Gaps 0;

QY 1 ANNHVSIWEAMSDNDTPYLH 20
Db 74 SNKHINILEDWMHNDPPEYFN 93

RESULT 8
US-09-995-749A-2
; Sequence 2, Application US/09995749A
; Patent No. US20020155568A1
; GENERAL INFORMATION:
; APPLICANT: VAN GEEL-SCHUTTEN, GERRITDINA HENDRIKA
; APPLICANT: DIJKHUIZEN, LUBBERT
; APPLICANT: RAHAOUI, HAKIM
; APPLICANT: LEER, ROBERT-JAN
; TITLE OF INVENTION: NOVEL GLUCOSYLTRANSFERASES
; FILE REFERENCE: BO43388-CIP
; CURRENT APPLICATION NUMBER: US/09/995,749A
; CURRENT FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: 09/604,957
; PRIOR FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: EPO 00201871.1
; PRIOR FILING DATE: 2000-05-25
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 2
; LENGTH: 1781
; TYPE: PRT
; ORGANISM: Lactobacillus reuteri
US-09-995-749A-2

Query Match 40.3%; Score 48; DB 10; Length 1781;
Best Local Similarity 35.0%; Pred. No. 1.6e+02;
Matches 7; Conservative 6; Mismatches 7; Indels 0; Gaps 0;

QY 1 ANNHVSIWEAMSDNDTPYLH 20
Db 1053 SNKHINILEDWMHNDPPEYFN 1072

RESULT 9
US-10-211-962-96
; Sequence 96, Application US/10211962
; Publication No. US20030082640A1
; GENERAL INFORMATION:
; APPLICANT: Hertz, Joachim
; APPLICANT: Gotthardt, Michael
; TITLE OF INVENTION: LDL Receptor Signaling Pathways
; FILE REFERENCE: UTSW0708
; CURRENT APPLICATION NUMBER: US/10/211,962
; CURRENT FILING DATE: 2002-08-01
; PRIOR APPLICATION NUMBER: US/09/562,737
; PRIOR FILING DATE: 2000-05-01
; NUMBER OF SEQ ID NOS: 132
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 96
; LENGTH: 749
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: Sequence
US-10-211-962-96

Query Match 39.5%; Score 47; DB 15; Length 749;
Best Local Similarity 53.3%; Pred. No. 86;
Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 6 STVAMSDNDTPYLH 20
Db 136 SAVEWTDSATPHPH 150

RESULT 10
US-09-888-615-73
; Sequence 73, Application US/09888615
; Patent No. US20020064856A1

GENERAL INFORMATION:
APPLICANT: PLOWMAN, GREGORY
APPLICANT: WHITE, DAVID
APPLICANT: CAENEPEEL, SEAN
APPLICANT: CHARVOCZAK, GLEN
APPLICANT: MANNING, GERARD
APPLICANT: SUDARSANAM, SUCHA
TITLE OF INVENTION: NOVEL PROTEASES
FILE REFERENCE: 038602/1214
CURRENT APPLICATION NUMBER: US/09/888,615
CURRENT FILING DATE: 2001-06-26
PRIOR APPLICATION NUMBER: 60/214,047
PRIOR FILING DATE: 2000-06-26
NUMBER OF SEQ ID NOS: 150
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 73
LENGTH: 1604
TYPE: PRT
ORGANISM: Homo sapiens
US-09-888-615-73

Query Match 39.1%; Score 46.5; DB 9; Length 1604;
Best Local Similarity 47.4%; Pred. No. 2.4e+02;
Matches 9; Conservative 3; Mismatches 4; Indels 3; Gaps 1;

QY 5 VSIVEAMSDN--DTPYH 20
DB 294 VALLEWKMDNRTDDEPLH 312

RESULT 11
US-10-009-823A-1
Sequence 1, Application US/10009823A
Publication No. US20030157120A1

GENERAL INFORMATION:
APPLICANT: Panaccio, Michael
APPLICANT: Rosey, Everett Lee
APPLICANT: Sinistaj, Meri
APPLICANT: Hasse, Declaf
APPLICANT: Parsons, Jim
APPLICANT: Ankenbauer, Robert G.
TITLE OF INVENTION: LAWSONIA DERIVED GENE AND RELATED FLGE
TITLE OF INVENTION: POLYPEPTIDES, PEPTIDES AND PROTEINS AND THEIR USES
FILE REFERENCE: DAV150.001APC
CURRENT APPLICATION NUMBER: US/10/009,823A
CURRENT FILING DATE: 2002-08-13
PRIOR APPLICATION NUMBER: PCT/AU00/00437
PRIOR FILING DATE: 2000-05-11
PRIOR APPLICATION NUMBER: US 60/133,973
PRIOR FILING DATE: 1999-05-13
NUMBER OF SEQ ID NOS: 13
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 1
LENGTH: 502
TYPE: PRT
ORGANISM: Lawsonia intracellularis
US-10-009-823A-1

Query Match 38.7%; Score 46; DB 12; Length 502;
Best Local Similarity 41.2%; Pred. No. 79;
Matches 7; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

QY 1 ANNHVSIVEAMSDNTP 17
DB 185 ANPYFALBSWKNGTIP 201

RESULT 12
US-10-195-144-17
Sequence 17, Application US/10195144
Publication No. US20030126646A1
GENERAL INFORMATION:
APPLICANT: BROWN, GREGORY G.

APPLICANT: FORMANOVA, NATASA
APPLICANT: DENDY, CHARLES
APPLICANT: LANDRY, BENOT S.
APPLICANT: CHEUNG, WING
APPLICANT: JIN, HUA
TITLE OF INVENTION: NUCLEAR FERTILITY RESTORER GENES AND METHODS OF USE IN
FILE REFERENCE: 16313-0136
CURRENT APPLICATION NUMBER: US/10/195,144
CURRENT FILING DATE: 2002-10-01
PRIOR APPLICATION NUMBER: 60/305,026
PRIOR FILING DATE: 2001-07-12
PRIOR APPLICATION NUMBER: 60/305,363
PRIOR FILING DATE: 2001-07-13
PRIOR APPLICATION NUMBER: 60/308,736
PRIOR FILING DATE: 2001-07-30
NUMBER OF SEQ ID NOS: 128
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 17
LENGTH: 1494
TYPE: PRT
ORGANISM: Raphanus sativum
US-10-195-144-17

Query Match 38.7%; Score 46; DB 15; Length 1494;
Best Local Similarity 41.2%; Pred. No. 2.7e+02;
Matches 7; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY 3 NHVSIVEAMSDNTPYL 19
DB 1130 NTSIAETWLNONSEPFL 1146

RESULT 13
US-09-738-626-4571
Sequence 4571, Application US/09738626
Publication No. US20020197605A1
GENERAL INFORMATION:
APPLICANT: NAKAGAWA, SATOSHI
APPLICANT: MIZOGUCHI, HIROSHI
APPLICANT: ANDO, SEIKO
APPLICANT: HAYASHI, MIKIRO
APPLICANT: OCHIAI, KEIKO
APPLICANT: YOKOI, HARUHIKO
APPLICANT: TATEISHI, NAOKO
APPLICANT: SENOH, AKIHIRO
APPLICANT: IKEDA, MASATO
APPLICANT: OZAKI, AKIO
TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
FILE REFERENCE: 249-125
CURRENT APPLICATION NUMBER: US/09/738,626
CURRENT FILING DATE: 2000-12-18
PRIOR APPLICATION NUMBER: JP 99/377484
PRIOR FILING DATE: 1999-12-16
PRIOR APPLICATION NUMBER: JP 00/159162
PRIOR FILING DATE: 2000-04-07
PRIOR APPLICATION NUMBER: JP 00/280988
PRIOR FILING DATE: 2000-08-03
NUMBER OF SEQ ID NOS: 7059
SOFTWARE: PatentIn Ver. 3.0
SEQ ID NO 4571
LENGTH: 199
TYPE: PRT
ORGANISM: Corynebacterium glutamicum
US-09-738-626-4571

Query Match 37.8%; Score 45; DB 10; Length 199;
Best Local Similarity 58.3%; Pred. No. 40;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 7 IVEAMSDNTPY 18
DB 177 LVEVMTDQDTRY 188

RESULT 14

US-10-156-761-14923
 ; Sequence 14923, Application US/10156761
 ; Publication No. US20030119018A1
 ; GENERAL INFORMATION:
 ; APPLICANT: OMIURA, SATOSHI
 ; APPLICANT: IKEDA, HARUO
 ; APPLICANT: ISHIKAWA, JUN
 ; APPLICANT: HORIKAWA, HIROSHI
 ; APPLICANT: SHIBA, TADAYOSHI
 ; APPLICANT: SAKAKI, YOSHIYUKI
 ; APPLICANT: HATTORI, MASAHIRA
 ; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
 ; FILE REFERENCE: 249-262
 ; CURRENT APPLICATION NUMBER: US/10/156,761
 ; CURRENT FILING DATE: 2002-05-29
 ; PRIOR APPLICATION NUMBER: JP 2001-204089
 ; PRIOR FILING DATE: 2001-05-30
 ; PRIOR APPLICATION NUMBER: JP 2001-272697
 ; PRIOR FILING DATE: 2001-08-02
 ; NUMBER OF SEQ ID NOS: 15109
 ; SEQ ID NO 14923
 ; LENGTH: 665
 ; TYPE: PRN
 ; ORGANISM: Streptomyces avermitilis
 US-10-156-761-14923

Query Match 37.8%; Score 45; DB 15; Length 665;
 Best Local Similarity 50.0%; Pred. No. 1.5e+02;
 Matches 9; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 3 NHVSIVEAMSDNDTPYLH 20
 DB 88 NWTYLVTEAMSDPTTWRH 105

RESULT 15
 US-09-924-154-16
 ; Sequence 16, Application US/09924154
 ; Patent No. US20020127241A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Nannum, David L.
 ; TITLE OF INVENTION: Anti-Plasmodium Compositions and Methods of Use
 ; FILE REFERENCE: 05213-0465 43170-262105
 ; CURRENT APPLICATION NUMBER: US/09/924,154
 ; CURRENT FILING DATE: 2001-08-07
 ; PRIOR APPLICATION NUMBER: US 60/223,525
 ; PRIOR FILING DATE: 2000-08-07
 ; NUMBER OF SEQ ID NOS: 17
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 16
 ; LENGTH: 972
 ; TYPE: PRN
 ; ORGANISM: Mammalian
 US-09-924-154-16

Query Match 37.8%; Score 45; DB 10; Length 972;
 Best Local Similarity 44.4%; Pred. No. 2.3e+02;
 Matches 8; Conservative 1; Mismatches 9; Indels 0; Gaps 0;

QY 3 NHVSIVEAMSDNDTPYLH 20
 DB 304 NHTGVCLMEDDDNDQYLH 321

Search completed: November 13, 2003, 10:29:04
 Job time : 28.4692 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Inc.

OM protein - protein search, using sw model

Run on: November 13, 2003, 09:31:40 ; Search time 40.3507 Seconds
(without alignments)
86.541 Million cell updates/sec

Title: US-09-290-049a-15
Perfect score: 110
Sequence: 1 VPSYFIRAHSEVQDLIRNII 22

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

A. Geneseq. 19Jun03:*

- 1: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
- 2: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
- 3: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
- 4: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
- 5: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
- 6: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
- 7: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
- 8: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
- 9: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
- 10: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
- 11: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
- 12: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
- 13: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
- 14: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
- 15: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
- 16: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
- 17: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
- 18: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
- 19: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
- 20: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
- 21: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
- 22: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*
- 23: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*
- 24: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	110	100.0	1375	AAU98028	S. mutans glucosyl
2	110	100.0	1375	AAU79288	Streptococcus muta
3	105	95.5	1017	AAU79285	Streptococcus muta
4	105	95.5	1476	AAU79284	Streptococcus muta
5	99	90.0	1475	AAU98027	S. mutans glucosyl
6	99	90.0	1475	AAU98030	S. mutans glucosyl
7	99	90.0	1475	AAU98031	S. mutans glucosyl
8	99	90.0	1475	AAU98033	S. mutans glucosyl
9	99	90.0	1475	AAU98039	S. mutans glucosyl

10	99	90.0	1475	23	AAU98040	S. mutans glucosyl
11	98	89.1	1475	23	AAU98035	S. mutans glucosyl
12	98	89.1	1592	14	AAU32925	Glucosyltransferase
13	92	83.6	1475	23	AAU98032	S. mutans glucosyl
14	92	83.6	1475	23	AAU98034	S. mutans glucosyl
15	91	82.7	1475	23	AAU98036	S. mutans glucosyl
16	91	82.7	1475	23	AAU98037	S. mutans glucosyl
17	91	82.7	1475	23	AAU98038	S. mutans GRP mut
18	82	74.5	1527	23	AAU80055	Leuconostoc mesent
19	78	70.9	2835	23	ABB98574	Dextran saccharase
20	78	69.1	15	23	ABB98570	Dextrane-saccharas
21	76	68.2	1781	23	AAU74519	Lactobacillus reut
22	75	68.2	1430	23	AAU98041	S. mutans glucosyl
23	75	68.2	1430	23	AAU98044	S. mutans glucosyl
24	71	64.5	15	23	ABB98651	Dextrane-saccharas
25	71	64.5	1430	23	AAU98042	S. mutans glucosyl
26	71	64.5	1430	23	AAU98045	S. mutans glucosyl
27	71	64.5	2057	21	AAU10667	L. mesenteroides a
28	70	63.6	1577	17	AAU91047	Alpha-D-glucosyltr
29	68	61.8	1430	23	AAU98029	S. mutans glucosyl
30	68	61.8	1430	23	AAU98043	S. mutans glucosyl
31	67	60.9	15	23	ABB98654	Dextrane-saccharas
32	65	59.1	15	23	ABB98653	Dextrane-saccharas
33	65	59.1	15	23	ABB98655	Dextrane-saccharas
34	63	57.3	15	23	ABB98657	Dextrane-saccharas
35	63	57.3	15	23	ABB98652	Dextrane-saccharas
36	57	51.8	15	23	ABB98656	Dextrane-saccharas
37	55	50.0	15	23	ABB98656	Dextrane-saccharas
38	51	46.4	855	23	ABB98573	Dextran saccharase
39	48	43.6	479	22	ABG20960	Novel human diago
40	47	42.7	541	8	AAU70433	Human atrial natr
41	47	42.7	541	17	AAU95586	Human atrial natr
42	47	42.7	639	20	AAU17089	Human atrial natr
43	45	40.9	1335	14	AAU42839	Human atrial natr
44	44.5	40.5	986	20	AAU37617	Urea amidolysase.
45	44	40.0	267	22	ABG24233	Protein which is s
						Novel human diago

ALIGNMENTS

RESULT 1	AAU98028	standard; Protein; 1375 AA.
XX	AAU98028;	
AC	27-AUG-2002	(first entry)
XX		
DE	S. mutans glucosyltransferase GTFC.	
XX		
KW	Glucosyltransferase; GTFC; transgenic plant; paper sizing;	
KM	coating composition; glucan; starch; latex; thermoplastic molecule;	
KW	amyloplastic; vacuole; paper manufacture.	
XX		
OS	Streptococcus mutans.	
XX		
EN	US2002031826-A1.	
XX		
PD	14-MAR-2002.	
XX		
PF	19-DEC-2000; 2000US-0740274.	
XX		
XX	11-DEC-1998; 98US-0210361.	
PR	07-JUN-1995; 95US-0478704.	
PR	07-JUN-1995; 95US-0482711.	
PR	07-JUN-1995; 95US-0485243.	
PR	16-JAN-1998; 98US-0007999.	
PR	16-JAN-1998; 98US-0008172.	
PR	20-JAN-1998; 98US-0009620.	
XX		
PA	(NICH//) NICHOLS S E.	
XX		

XX Nichols SE;
XX WPI, 2002-414332/44.
DR N-PSDB; ABKS2939.
XX
PT Glucosyltransferase B or D protein useful for producing a glucan useful
PT as substitutes for and additions to modified starch and latexes in
PT paper manufacture, comprises mutations in specific positions -
XX
XX Disclosure, Page 30-33; 44pp; English.
PS
XX
XX The invention an isolated protein comprising a glucosyltransferase
CC (GTF) B polypeptide having changes at position from I448V, D457N,
CC D567T, K1014T, D457N/D567T, D457N/D571K, D567T/D571K,
CC D567T/D571K/K1014T, I448V/D457N/D567T/D571K/K779Q/K1014T,
CC Y169A/Y170A/Y171A, and K779Q or a GTF D polypeptide having
CC changes at positions from T589D, T589E, M471D, M471D/T589D, and
CC M471D/T589E. Also included are a glucan produced by the GTF mutant,
CC an isolated polynucleotide which encodes P1 or P2, or its complementary
CC polynucleotide, a ribonucleic acid sequence encoding the GTF mutant,
CC an expression cassette comprising the polynucleotide operably linked to a
CC promoter, a vector comprising the expression cassette, host cell
CC introduced with the vector, a transgenic plant comprising the
CC vector, a seed or tuber from the transgenic plant, a paper sling and/or
CC coating composition comprising a glucan produced in a plant transformed
CC with a gene encoding the mutant GTF, wild type or, starch, a latex,
CC thermoplastic molecule or their combinations or glucan and starch where
CC the glucan is produced in the amyloplast and/or vacuole or a maize line
CC deficient in starch biosynthesis, transformed with a gene encoding a
CC glucosyltransferase B or D enzyme, wild-type or mutant and a paper
CC comprising the glucan (paper sizing/coating agent). The vector is useful
CC for producing a glucan in a plant. The method comprises transforming a
CC plant cell with the vector, growing the plant cell under plant growing
CC conditions to produce a regenerated plant and inducing expression of the
CC polynucleotide for a time sufficient to produce the glucan in the
CC regenerated plant, where the vector contains a transit sequence from
CC ribulose biphosphate carboxylase small subunit, waxy, brittle-1 and
CC chlorophyll AB binding protein to produce a transgenic plant, and glucan
CC is produced in the amyloplast of potato or the vacuole of sugar beet.
CC Glucans are useful as substitutes for and additions to modified starch
CC and latexes in paper manufacture. Unlike prior art techniques, which
CC require input materials that produce chemical effluents, paper
CC manufacture utilising the glucan produced by GTF, which utilises
CC environmentally produced input materials, is more cost-effective and
CC biologically friendly. Moreover, glucans also exhibit thermoplastic
CC properties and impart gloss to the paper during coating step.
CC The present sequence represents GTFCC.
XX
XX
SQ Sequence 1375 AA;
QY
DB 578 VPSYSFIRAHDSFVODLIRNI 599
1 VPSYSFIRAHDSFVODLIRNI 22
|||||
Query Match 100.0%; Score 110; DB 23; Length 1375;
Best Local Similarity 100.0%; Pred. No. 1.2e-09;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1 VPSYSFIRAHDSFVODLIRNI 22
|||||
AAU79288
ID AAU79288 standard; Protein; 1375 AA.
XX
XX AAU79288;
XX
XX 13-AUG-2002 (first entry)
XX
XX Streptococcus mutans monoclonal antibody-related protein #5.
DE
XX Antibody; dental caries; water insoluble glucan synthetase;
KM anti-caries; glucosyl transferase-B; immunotherapy.
XX
XX Streptococcus mutans.
OS

```

PN      JP2002114709-A.
XX
XX      16-APR-2002.
PD
XX      04-OCT-2000; 2000JP-0304889.
PF
XX      04-OCT-2000; 2000JP-0304889.
XX
XX      (UNNT-) UNIV NIPPON.
PA
XX      WPI; 2002-448101/48.
XX
XX      Anti-carries agent composed of a monoclonal antibody against an
PT      inhibitory enzyme against water insoluble glucan synthetase of glucosyl
PT      transferase-B (GTF-B) of Streptococcus mutans
PT
XX      Disclosure; Page 22-25; 28pp; Japanese.
PS
XX
XX      The invention relates to a monoclonal antibody against dental caries and
CC      an anti-carries agent composed of a monoclonal antibody produced by
CC      Streptococcus mutans, particularly mouse-hybridoma MHP126 (FERM P-17566)
CC      or mouse-hybridoma MHP136 (FERM P-17567), against an enzyme having
CC      inhibitive activity against water insoluble glucan synthetase of glucosyl
CC      transferase-B. The monoclonal antibody specifically inhibits water
CC      insoluble glucan synthetase of Streptococcus mutans produced glucosyl
CC      transferase-B and is used in the immunotherapy of dental caries. This
CC      sequence represents a Streptococcus mutans monoclonal antibody-related
CC      protein.
CC
XX      Sequence      1375 AA;
SQ
XX
XX      Query Match      100.0%; Score 110; DB 23; Length 1375;
XX      Best Local Similarity 100.0%; Pred. No. 1.2e-09;
XX      Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0
XX
XX      1 VPSYSFIRAHDSVEVDLIRNTI 22
XX      |||||
XX      578 VPSYSFIRAHDSVEVDLIRNTI 599
XX
XX      RESULT 3
XX      ID AAU79285
XX      AAU79285 standard. Protein; 1017 AA.
XX
XX      AAU79285;
XX
XX      13-AUG-2002 (first entry)
XX
XX      Streptococcus mutans monoclonal antibody-related protein #2.
XX
XX      Antibody; dental caries; water insoluble glucan synthetase;
XX      anti-carries; glucosyl transferase-B; immunotherapy.
XX
XX      Streptococcus mutans.
XX
XX      JP2002114709-A.
XX
XX      16-APR-2002.
XX
XX      04-OCT-2000; 2000JP-0304889.
XX
XX      04-OCT-2000; 2000JP-0304889.
XX
XX      (UNNT-) UNIV NIPPON.
XX
XX      WPI; 2002-448101/48.
XX
XX      Anti-carries agent composed of a monoclonal antibody against an
PT      inhibitory enzyme against water insoluble glucan synthetase of glucosyl
PT      transferase-B (GTF-B) of Streptococcus mutans
PT
XX      Claim 4; Page 17-19; 28pp; Japanese.

```

XX The invention relates to a monoclonal antibody against dental caries and
 CC an anti-carries agent composed of a monoclonal antibody produced by
 CC Streptococcus mutans, particularly mouse-hybridoma MHP126 (FERM P-17566)
 CC or mouse-hybridoma MHP136 (FERM P-17567), against an enzyme having
 CC inhibitive activity against water insoluble glucan synthetase of glucosyl
 CC transferase-B. The monoclonal antibody specifically inhibits water
 CC insoluble glucan synthetase of Streptococcus mutans produced glucosyl
 CC transferase-B and is used in the immunotherapy of dental caries. This
 CC sequence represents a Streptococcus mutans monoclonal antibody-related
 CC protein.
 XX
 SO Sequence 1017 AA;
 Query Match 95.5%; Score 105; DB 23; Length 1017;
 Best Local Similarity 95.5%; Pred. No. 6.2e-09;
 Matches 21; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 1 VPSYSFIRAHDSFVQDLIRNII 22
 518 VPSYSFIRAHDSFVQDLIRDII 539
 DB
 RESULT 4
 AAU79284
 ID AAU79284 standard; Protein; 1476 AA.
 AC AAU79284;
 XX
 XX 13-AUG-2002 (first entry)
 DT
 XX Streptococcus mutans monoclonal antibody-related protein #1.
 DE Streptococcus mutans monoclonal antibody-related protein #1.
 XX Antibody; dental caries; water insoluble glucan synthetase;
 KM anti-carries; glucosyl transferase-B; immunotherapy.
 XX Streptococcus mutans.
 OS
 XX JP2002114709-A.
 PN 16-APR-2002.
 PD
 XX 04-OCT-2000; 2000JP-0304889.
 PF 04-OCT-2000; 2000JP-0304889.
 XX 04-OCT-2000; 2000JP-0304889.
 PR
 XX (UYN1-) UNIV NIPPON.
 PA
 XX WPI; 2002-448101/48.
 DR
 XX Anti-carries agent composed of a monoclonal antibody against an
 PT inhibitory enzyme against water insoluble glucan synthetase of glucosyl
 PT transferase-B (GTF-B) of Streptococcus mutans
 XX
 XX Claim 3; Page 13-16; 28pp; Japanese.
 PS
 XX The invention relates to a monoclonal antibody against dental caries and
 CC an anti-carries agent composed of a monoclonal antibody produced by
 CC Streptococcus mutans, particularly mouse-hybridoma MHP126 (FERM P-17566)
 CC or mouse-hybridoma MHP136 (FERM P-17567), against an enzyme having
 CC inhibitive activity against water insoluble glucan synthetase of glucosyl
 CC transferase-B. The monoclonal antibody specifically inhibits water
 CC insoluble glucan synthetase of Streptococcus mutans produced glucosyl
 CC transferase-B and is used in the immunotherapy of dental caries. This
 CC sequence represents a Streptococcus mutans monoclonal antibody-related
 CC protein.
 XX
 SO Sequence 1476 AA;
 Query Match 95.5%; Score 105; DB 23; Length 1476;
 Best Local Similarity 95.5%; Pred. No. 9.6e-09;
 Matches 21; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 VPSYSFIRAHDSFVQDLIRNII 22
 DB 552 VPSYSFIRAHDSFVQDLIRDII 573
 RESULT 5
 AAU98027
 ID AAU98027 standard; Protein; 1475 AA.
 XX
 XX AAU98027;
 AC
 XX 27-AUG-2002 (first entry)
 DT
 XX S. mutans glucosyltransferase GTFB.
 DE
 XX Glucosyltransferase; GTFB; transgenic plant; paper sizing;
 KM coating composition; glucan; starch; latex; thermoplastic molecule;
 KM amyloplast; vacuole; paper manufacture.
 XX Streptococcus mutans.
 OS
 XX US2002031826-A1.
 PN
 XX 14-MAR-2002.
 PD
 XX 19-DEC-2000; 2000US-0740274.
 PF
 XX 11-DEC-1998; 98US-0210361.
 PR 07-JUN-1995; 95US-0478704.
 PR 07-JUN-1995; 95US-0482711.
 PR 07-JUN-1995; 95US-0485243.
 PR 16-JAN-1998; 98US-0007999.
 PR 16-JAN-1998; 98US-0008172.
 PR 20-JAN-1998; 98US-0009620.
 XX (NICH/) NICHOLS S E.
 PA
 XX Nichols SE;
 PI
 XX WPI; 2002-414332/44.
 DR N-PADB; ABK52938.
 XX
 XX Glucosyltransferase B or D protein useful for producing a glucan useful
 PT as substitutes for and additions to modified starch and latexes in
 PT paper manufacture, comprises mutations in specific positions
 XX
 PS Disclosure; Page 21-25; 44pp; English.
 XX
 XX The invention an isolated protein comprising a glucosyltransferase
 CC (GTF) B polypeptide having changes at position from I448V, D457N,
 CC D567T, K1014T, D457N/D571K, D457N/D571K, D567T/D571K,
 CC D567T/D571K/K1014T, I448V/D457N/D567T/D571K/K779G/K1014T,
 CC Y169A/Y170A/Y171A, and K779G or a GTF D polypeptide having
 CC changes at positions from T589D, T589E, N471D, N471D/T589D, and
 CC N471D/T589E. Also included are a glucan produced by the GTF mutant,
 CC an isolated polynucleotide which encodes FI or P2, or its complementary
 CC an expression cassette comprising the polynucleotide operably linked to a
 CC promoter, a vector comprising the expression cassette, host cell
 CC introduced with the vector, a transgenic plant comprising the
 CC vector, a seed or tuber from the transgenic plant, a paper sizing and/or
 CC coating composition comprising a glucan produced in a plant transformed
 CC with a gene encoding the mutant GTF, wild type or, starch, a latex,
 CC thermoplastic molecule or their combinations or glucan and starch where
 CC the glucan is produced in the amyloplast and/or vacuole or a maize line
 CC deficient in starch biosynthesis, transformed with a gene encoding a
 CC glucosyltransferase B or D enzyme, wild-type or mutant and a paper
 CC comprising the glucan (paper sizing/coating agent). The vector is useful
 CC for producing a glucan in a plant. The method comprises transforming a
 CC plant cell with the vector, growing the plant cell under plant growing
 CC conditions to produce a regenerated plant and inducing expression of the
 CC polynucleotide for a time sufficient to produce the glucan in the
 CC regenerated plant, where the vector contains a transit sequence from

CC ribulose biphosphate carboxylase small subunit, waxy, brittle-1 and
 CC chlorophyll AB binding protein to produce a transgenic plant, and glucan
 CC is produced in the amyloplast of potato or the vacuole of sugar beet.
 CC Glucans are useful as substitutes for and additions to modified starch
 CC and latexes in paper manufacture. Unlike prior art techniques, which
 CC require input materials that produce chemical effluents, paper
 CC manufacture utilizing the glucan produced by GTF, which utilizes
 CC biologically produced input materials, is more cost-effective and
 CC environmentally friendly. Moreover, glucans also exhibit thermoplastic
 CC properties and impart gloss to the paper during coating step.
 CC The present sequence represents GTFB.
 CC
 XX
 XX Sequence 1475 AA;
 SQ
 Query Match 90.0%; Score 99; DB 23; Length 1475;
 Best Local Similarity 90.9%; Pred. No. 1e-07;
 Matches 20; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 QY 1 VPSYSFIRAHDSVEVDIIRNII 22
 DB 552 VPSYSFIRAHDSVEVDIIRNII 573
 RESULT 6
 AAU98030 ID AAU98030 standard; Protein; 1475 AA.
 XX
 XX AAU98030;
 AC
 XX 27-AUG-2002 (first entry)
 DT
 XX
 XX S. mutans glucosyltransferase GTFB mutant I448V.
 DE
 XX Glucosyltransferase; GTFB; transgenic plant; paper sizing;
 KM coating composition; glucan; starch; latex; thermoplastic molecule;
 KM amyloplast; vacuole; paper manufacture; mutant; mutain.
 KM
 XX Streptococcus mutans.
 OS
 XX Synthetic.
 OS
 XX Key Location/Qualifiers
 FH Misc-difference 448
 FT /note= "Wild-type Ile substituted by Val"
 FT
 XX US2002031826-A1.
 PN
 XX
 XX 14-MAR-2002.
 PD
 XX
 XX 19-DEC-2000; 2000US-0740274.
 PF
 XX 11-DEC-1998; 98US-0210361.
 PR 07-JUN-1995; 95US-0478704.
 PR 07-JUN-1995; 95US-0482711.
 PR 07-JUN-1995; 95US-0485243.
 PR 16-JAN-1998; 98US-0007939.
 PR 16-JAN-1998; 98US-0008172.
 PR 20-JAN-1998; 98US-0009620.
 PR
 XX (NICH/) NICHOLS S E.
 PA
 XX Nichols SE;
 PI
 XX WPI; 2002-414332/44.
 DR
 XX Glucosyltransferase B or D protein useful for producing a glucan useful
 PT as substitutes for and additions to modified starch and latexes in
 PT paper manufacture, comprises mutations in specific positions -
 XX
 XX Claim 36; Page -; 44pp; English.
 PS
 XX The invention an isolated protein comprising a glucosyltransferase
 CC (GTF) B polypeptide having changes at position from I448V, D457N,
 CC D567T, K1014T, D457N/D567T, D457N/D571K, D567T/D571K,

CC D567T/D571K/K1014T, I448V/D457N/D567T/D571K/K779Q/K1014T.
 CC Y169A/Y170A/Y171A, and K779Q or a GTF D polypeptide having
 CC changes at positions from T589D, T589E, N471D, N471D/T589D, and
 CC N471D/T589E. Also included are a glucan produced by the GTF mutant,
 CC an isolated polynucleotide which encodes P1 or P2, or its complementary
 CC polynucleotide, a ribonucleic acid sequence encoding the GTF mutant,
 CC an expression cassette comprising the polynucleotide operably linked to a
 CC promoter, a vector comprising the expression cassette, host cell
 CC introduced with the vector, a transgenic plant comprising the
 CC vector, a seed or tuber from the transgenic plant, a paper sizing and/or
 CC coating composition comprising a glucan produced in a plant transformed
 CC with a gene encoding the mutant GTF, wild type or, starch, a latex,
 CC thermoplastic molecule or their combinations or glucan and starch where
 CC the glucan is produced in the amyloplast and/or vacuole of a maize line
 CC deficient in starch biosynthesis, transformed with a gene encoding a
 CC glucosyltransferase B or D enzyme, wild-type or mutant and a paper
 CC comprising the glucan (paper sizing/coating agent). The vector is useful
 CC for producing a glucan in a plant. The method comprises transforming a
 CC plant cell with the vector, growing the plant cell under plant growing
 CC conditions to produce a regenerated plant and inducing expression of the
 CC polynucleotide for a time sufficient to produce the glucan in the
 CC regenerated plant, where the vector contains a transit sequence from
 CC ribulose biphosphate carboxylase small subunit, waxy, brittle-1 and
 CC chlorophyll AB binding protein to produce a transgenic plant, and glucan
 CC is produced in the amyloplast of potato or the vacuole of sugar beet.
 CC Glucans are useful as substitutes for and additions to modified starch
 CC and latexes in paper manufacture. Unlike prior art techniques, which
 CC require input materials that produce chemical effluents, paper
 CC manufacture utilizing the glucan produced by GTF, which utilizes
 CC biologically produced input materials, is more cost-effective and
 CC environmentally friendly. Moreover, glucans also exhibit thermoplastic
 CC properties and impart gloss to the paper during coating step.
 CC The present sequence represents a GTFB mutant of the invention.
 CC Note: The present sequence is not shown in the specification but
 CC was created by the indexer using the GTFB sequence appearing as AAU98027
 CC and the information in claim 36.
 CC
 XX
 XX Sequence 1475 AA;
 SQ
 Query Match 90.0%; Score 99; DB 23; Length 1475;
 Best Local Similarity 90.9%; Pred. No. 1e-07;
 Matches 20; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 QY 1 VPSYSFIRAHDSVEVDIIRNII 22
 DB 552 VPSYSFIRAHDSVEVDIIRNII 573
 RESULT 7
 AAU98031 ID AAU98031 standard; Protein; 1475 AA.
 XX
 XX AAU98031;
 AC
 XX 27-AUG-2002 (first entry)
 DT
 XX S. mutans glucosyltransferase GTFB mutant D457N.
 DE
 XX Glucosyltransferase; GTFB; transgenic plant; paper sizing;
 KM coating composition; glucan; starch; latex; thermoplastic molecule;
 KM amyloplast; vacuole; paper manufacture; mutant; mutain.
 KM
 XX Streptococcus mutans.
 OS
 XX Synthetic.
 OS
 XX Key Location/Qualifiers
 FH Misc-difference 457
 FT /note= "Wild-type Asp substituted by Asn"
 FT
 XX US2002031826-A1.
 PN
 XX
 XX 14-MAR-2002.
 PD
 XX

PF 19-DEC-2000; 2000US-0740274.
 XX 11-DEC-1998; 98US-0210361.
 PR 07-JUN-1995; 95US-0478704.
 PR 07-JUN-1995; 95US-0482711.
 PR 07-JUN-1995; 95US-0485243.
 PR 16-JAN-1998; 98US-0007999.
 PR 16-JAN-1998; 98US-0008172.
 PR 20-JAN-1998; 98US-0009620.
 PA (NICH/) NICHOLS S E.
 XX
 PI Nichols SE;
 XX WPI; 2002-414332/44.
 DR
 XX Glucosyltransferase B or D protein useful for producing a glucan useful
 PT as substitutes for and additions to modified starch and latexes in
 PT paper manufacture, comprises mutations in specific positions -
 XX
 PS Claim 36; Page -; 44pp; English.

CC The invention an isolated protein comprising a glucosyltransferase
 CC (GTF) B polypeptide having changes at position from 1448V, D457N,
 CC D567T, K1014T, D457N/D567T, D457N/D571K, D567T/D571K,
 CC D567T/D571K/K1014T, 1448V/D457N/D567T/K779Q/K1014T,
 CC Y169A/Y170A/Y171A, and K779Q or a GTF D polypeptide having
 CC changes at positions from T589D, T589E, N471D, N471D/T589D and
 CC N471D/T589E. Also included are a glucan produced by the GTF mutant,
 CC an isolated polynucleotide which encodes P1 or P2, or its complementary
 CC polynucleotide, a ribonucleic acid sequence encoding the GTF mutant,
 CC an expression cassette comprising the polynucleotide operably linked to a
 CC promoter, a vector comprising the expression cassette, host cell
 CC introduced with the vector, a transgenic plant comprising the
 CC vector, a seed or tuber from the transgenic plant, a paper sizing and/or
 CC coating composition comprising a glucan produced in a plant transformed
 CC with a gene encoding the mutant GTF, wild type or, starch, a latex,
 CC thermoplastic molecule or their combinations or glucan and starch where
 CC the glucan is produced in the amyloplast and/or vacuole or a maize line
 CC deficient in starch biosynthesis, transformed with a gene encoding a
 CC glucosyltransferase B or D enzyme, wild-type or mutant and a paper
 CC comprising the glucan (paper sizing/coating agent). The vector is useful
 CC for producing a glucan in a plant. The method comprises transforming a
 CC plant cell with the vector, growing the plant cell under plant growing
 CC conditions to produce a regenerated plant and inducing expression of the
 CC polynucleotide for a time sufficient to produce the glucan in the
 CC regenerated plant, where the vector contains a transit sequence from
 CC ribulose biphosphate carboxylase small subunit, waxy, brittle-1 and
 CC chlorophyll AB binding protein to produce a transgenic plant, and glucan
 CC is produced in the amyloplast of potato or the vacuole of sugar beet.
 CC Glucans are useful as substitutes for and additions to modified starch
 CC and latexes in paper manufacture. Unlike prior art techniques, which
 CC require input materials that produce chemical effluents, paper
 CC manufacture utilizing the glucan produced by GTF, which utilizes
 CC biologically produced input materials, is more cost-effective and
 CC environmentally friendly. Moreover, glucans also exhibit thermoplastic
 CC properties and impart gloss to the paper during coating step.
 CC The present sequence represents a GTF mutant of the invention.
 CC Note: The present sequence is not shown in the specification but
 CC was created by the indexer using the GTFB sequence appearing as AAU98027
 CC and the information in claim 36.

XX Sequence 1475 AA;

Query Match 90.0%; Score 99; DB 23; Length 1475;

Best Local Similarity 90.9%; Pred. No. 1e-07; 1; Indels 0; Gaps 0;
 Matches 20; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 VPSYSEIRAHSEVODLLRNTI 22
 DB 552 VPSYSEIRAHSEVODLLRNTI 573

RESULT 8
 AAU98033
 ID AAU98033 standard; Protein; 1475 AA.
 XX
 AC AAU98033;
 XX
 DT 27-AUG-2002 (first entry)
 XX
 DE S. mutans glucosyltransferase GTFB mutant K1014T.
 XX
 KM Glucosyltransferase; GTFB; transgenic plant; paper sizing;
 KM coating composition; glucan; starch; latex; thermoplastic molecule;
 KM amyloplast; vacuole; paper manufacture; mutant; mutain.
 XX
 OS Streptococcus mutans.
 OS Synthetic.
 XX
 FH Key Location/Qualifiers
 FT Misc-difference 1014
 FT /note= "wild-type Lys substituted by Thr"
 XX
 XX US2002031826-A1.
 XX
 PD 14-MAR-2002.
 XX
 PF 19-DEC-2000; 2000US-0740274.
 XX
 XX 11-DEC-1998; 98US-0210361.
 PR 07-JUN-1995; 95US-0478704.
 PR 07-JUN-1995; 95US-0482711.
 PR 07-JUN-1995; 95US-0485243.
 PR 16-JAN-1998; 98US-0007999.
 PR 16-JAN-1998; 98US-0008172.
 PR 20-JAN-1998; 98US-0009620.
 XX
 PA (NICH/) NICHOLS S E.
 XX
 PI Nichols SE;
 XX WPI; 2002-414332/44.
 DR
 XX Glucosyltransferase B or D protein useful for producing a glucan useful
 PT as substitutes for and additions to modified starch and latexes in
 PT paper manufacture, comprises mutations in specific positions -
 XX
 PS Claim 36; Page -; 44pp; English.

CC The invention an isolated protein comprising a glucosyltransferase
 CC (GTF) B polypeptide having changes at position from 1448V, D457N,
 CC D567T, K1014T, D457N/D567T, D457N/D571K, D567T/D571K,
 CC D567T/D571K/K1014T, 1448V/D457N/D567T/K779Q/K1014T,
 CC Y169A/Y170A/Y171A, and K779Q or a GTF D polypeptide having
 CC changes at positions from T589D, T589E, N471D, N471D/T589D and
 CC N471D/T589E. Also included are a glucan produced by the GTF mutant,
 CC an isolated polynucleotide which encodes P1 or P2, or its complementary
 CC polynucleotide, a ribonucleic acid sequence encoding the GTF mutant,
 CC an expression cassette comprising the polynucleotide operably linked to a
 CC promoter, a vector comprising the expression cassette, host cell
 CC introduced with the vector, a transgenic plant comprising the
 CC vector, a seed or tuber from the transgenic plant, a paper sizing and/or
 CC coating composition comprising a glucan produced in a plant transformed
 CC with a gene encoding the mutant GTF, wild type or, starch, a latex,
 CC thermoplastic molecule or their combinations or glucan and starch where
 CC the glucan is produced in the amyloplast and/or vacuole or a maize line
 CC deficient in starch biosynthesis, transformed with a gene encoding a
 CC glucosyltransferase B or D enzyme, wild-type or mutant and a paper
 CC comprising the glucan (paper sizing/coating agent). The vector is useful
 CC for producing a glucan in a plant. The method comprises transforming a
 CC plant cell with the vector, growing the plant cell under plant growing
 CC conditions to produce a regenerated plant and inducing expression of the
 CC polynucleotide for a time sufficient to produce the glucan in the
 CC regenerated plant, where the vector contains a transit sequence from
 CC ribulose biphosphate carboxylase small subunit, waxy, brittle-1 and

CC chlorophyll AB binding protein to produce a transgenic plant, and glucan
 CC is produced in the amyloplast of potato or the vacuole of sugar beet.
 CC Glucans are useful as substitutes for and additions to modified starch
 CC and latexes in paper manufacture. Unlike prior art techniques, which
 CC require input materials that produce chemical effluents, paper
 CC manufacture utilizing the glucan produced by GTF, which utilizes
 CC biologically produced input materials, is more cost-effective and
 CC environmentally friendly. Moreover, glucans also exhibit thermoplastic
 CC properties and impart gloss to the paper during coating step.
 CC The present sequence represents a GTFB mutant of the invention.
 CC Note: The present sequence is not shown in the specification but
 CC was created by the indexer using the GTFB sequence appearing as AAU98027
 CC and the information in claim 36.

CC Sequence 1475 AA;

Query Match 90.0%; Score 99; DB 23; Length 1475;
 Best Local Similarity 90.9%; Pred. No. 1e-07;
 Matches 20; Conservative 1; Mismatches 0; Gaps 0;

Qy 1 VPSYSFIRAHDSFVQDLIRNII 22
 Db 552 VPSYSFIRAHDSFVQDLIRNII 573

RESULT 9
 AAU98039
 ID AAU98039 standard; Protein; 1475 AA.

AC AAU98039;

DT 27-AUG-2002 (first entry)

DE S. mutans glucosyltransferase GTFB mutant YYY169-171AA.

XX Glucosyltransferase; GTFB; transgenic plant; paper sizing;
 XX coating composition; glucan; starch; latex; thermoplastic molecule;

KM amyloplast; vacuole; paper manufacture; mutant; mutein.

OS Streptococcus mutans.

OS Synthetic.

XX Key Location/Qualifiers

FT Misc-difference 169..171
 FT /note= "Wild-type Tyr-Tyr-Tyr substituted by
 FT Ala-Ala-Ala"

XX US2002031826-A1.

XX 14-MAR-2002.

XX 19-DEC-2000; 2000US-0740274.

XX 11-DEC-1998; 98US-0210361.

XX 07-JUN-1995; 95US-0487704.

XX 07-JUN-1995; 95US-0482711.

XX 07-JUN-1995; 95US-0485243.

XX 16-JAN-1998; 98US-0007999.

XX 16-JAN-1998; 98US-0008172.

XX 20-JAN-1998; 98US-0009620.

XX (NICH/) NICHOLS S E.

XX Nichols SE;

XX WPI; 2002-414332/44.

XX Glucosyltransferase B or D protein useful for producing a glucan useful
 XX as substitutes for and additions to modified starch and latexes in
 XX paper manufacture, comprises mutations in specific positions
 XX Claim 36; Page -; 44pp; English.

CC The invention an isolated protein comprising a glucosyltransferase
 CC (GTF) B polypeptide having changes at position from 148V, D457N,
 CC D567T, K1014T, D457N/D567T, D457N/D571K, D567T/D571K,
 CC D567T/D571K/K1014T, 1448V/D457N/D567T/D571K/K790Q/K1014T,
 CC Y168A/Y170A/Y171A, and K790Q or a GTF D polypeptide having
 CC changes at positions from 1589D, T589E, N471D, N471D/Y589D, and
 CC N471D/Y589E. Also included are a glucan produced by the GTF mutant,
 CC an isolated polynucleotide which encodes P1 or P2, or its complementary
 CC polynucleotide, a ribonucleic acid sequence encoding the GTF mutant,
 CC an expression cassette comprising the polynucleotide operably linked to a
 CC promoter, a vector comprising the expression cassette, host cell
 CC introduced with the vector, a transgenic plant comprising the
 CC vector, a seed or tuber from the transgenic plant in a plant transformed
 CC coating composition comprising a glucan produced in a plant transformed
 CC with a gene encoding the mutant GTF, wild type or, starch, a latex,
 CC thermoplastic molecule or their combinations and/or vacuole or a maize line
 CC the glucan is produced in the amyloplast and/or vacuole of a plant
 CC deficient in starch biosynthesis, transformed with a gene encoding a
 CC glucosyltransferase B or D enzyme, wild-type or mutant and a paper
 CC comprising the glucan (paper sizing/coating agent). The vector is useful
 CC for producing a glucan in a plant. The method comprises transforming a
 CC plant cell with the vector, growing the plant cell under plant growing
 CC conditions to produce a regenerated plant and inducing expression of the
 CC polynucleotide for a time sufficient to produce the glucan in the
 CC regenerated plant, where the vector contains a transit sequence from
 CC ribulose biphosphate carboxylase small subunit, waxy, brittle-1 and
 CC chlorophyll AB binding protein to produce a transgenic plant, and glucan
 CC is produced in the amyloplast of potato or the vacuole of sugar beet.
 CC Glucans are useful as substitutes for and additions to modified starch
 CC and latexes in paper manufacture. Unlike prior art techniques, which
 CC require input materials that produce chemical effluents, paper
 CC manufacture utilizing the glucan produced by GTF, which utilizes
 CC biologically produced input materials, is more cost-effective and
 CC environmentally friendly. Moreover, glucans also exhibit thermoplastic
 CC properties and impart gloss to the paper during coating step.
 CC The present sequence represents a GTFB mutant of the invention.
 CC Note: The present sequence is not shown in the specification but
 CC was created by the indexer using the GTFB sequence appearing as AAU98027
 CC and the information in claim 36.

CC Sequence 1475 AA;

Query Match 90.0%; Score 99; DB 23; Length 1475;
 Best Local Similarity 90.9%; Pred. No. 1e-07;
 Matches 20; Conservative 1; Mismatches 0; Gaps 0;

Qy 1 VPSYSFIRAHDSFVQDLIRNII 22
 Db 552 VPSYSFIRAHDSFVQDLIRNII 573

RESULT 10
 AAU98040
 ID AAU98040 standard; Protein; 1475 AA.

AC AAU98040;

DT 27-AUG-2002 (first entry)

DE S. mutans glucosyltransferase GTFB mutant K779Q.

XX Glucosyltransferase; GTFB; transgenic plant; paper sizing;
 XX coating composition; glucan; starch; latex; thermoplastic molecule;

KM amyloplast; vacuole; paper manufacture; mutant; mutein.

OS Streptococcus mutans.

OS Synthetic.

XX Key Location/Qualifiers

FT Misc-difference 779
 FT /note= "Wild-type Lys substituted by Gln"

XX US2002031826-A1.

XX 14-MAR-2002.
 PD 19-DEC-2000; 2000US-0740274.
 PF 11-DEC-1998; 98US-0210361.
 PR 07-JUN-1995; 95US-0478704.
 PR 07-JUN-1995; 95US-0482711.
 PR 07-JUN-1995; 95US-0485243.
 PR 16-JAN-1998; 98US-0007999.
 PR 16-JAN-1998; 98US-0008172.
 PR 20-JAN-1998; 98US-0009620.
 XX (NICH/) NICHOLS S E.
 PA Nichols SE;
 PI WPI; 2002-414332/44.
 DR Glucosyltransferase B or D protein useful for producing a glucan useful
 PT as substitutes for and additions to modified starch and latexes in
 PT paper manufacture, comprises mutations in specific positions -
 XX
 XX Claim 36; Page -; 44pp; English.
 PS
 XX The invention an isolated protein comprising a glucosyltransferase
 CC (GTF) B polypeptide having changes at position from 1448V, D457N,
 CC D567T, K1014T, D457N/D567T, D457N/D571K, D567T/D571K,
 CC D567T/D571K/K1014T, 1448V/D457N/D567T/D571K/K779Q/K1014T,
 CC Y169A/Y170A/Y171A, and K779Q or a GTF D polypeptide having
 CC changes at positions from T589D, T589E, N471D, N471D/T589D, and
 CC N471D/T589E. Also included are a glucan produced by the GTF mutant,
 CC an isolated polynucleotide which encodes P1 or P2, or its complementary
 CC polynucleotide, a ribonucleic acid sequence encoding the GTF mutant,
 CC an expression cassette comprising the polynucleotide operably linked to a
 CC promoter, a vector comprising the expression cassette, host cell
 CC introduced with the vector, a transgenic plant comprising the
 CC vector, a seed or tuber from the transgenic plant, a paper string and/or
 CC coating composition comprising a glucan produced in a plant transformed
 CC with a gene encoding the mutant GTF, wild type or, starch, a latex,
 CC thermoplastic molecule or their combinations or glucan and starch where
 CC the glucan is produced in the amyloplast and/or vacuole or a maize line
 CC deficient in starch biosynthesis, transformed with a gene encoding a
 CC glucosyltransferase B or D enzyme, wild-type or mutant and a paper
 CC comprising the glucan (paper sizing/coating agent). The vector is useful
 CC for producing a glucan in a plant. The method comprises transforming a
 CC plant cell with the vector, growing the plant cell under plant growing
 CC conditions to produce a regenerated plant and inducing expression of the
 CC polynucleotide for a time sufficient to produce the glucan in the
 CC regenerated plant, where the vector contains a transit sequence from
 CC ribulose biphosphate carboxylase small subunit, waxy, brittle-1 and
 CC chlorophyll AB binding protein to produce a transgenic plant, and glucan
 CC is produced in the amyloplast of potato or the vacuole of sugar beet.
 CC Glucans are useful as substitutes for and additions to modified starch
 CC and latexes in paper manufacture. Unlike prior art techniques, which
 CC require input materials that produce chemical effluents, paper
 CC manufacture utilizing the glucan produced by GTF, which utilizes
 CC biologically produced input materials, is more cost-effective and
 CC environmentally friendly. Moreover, glucans also exhibit thermoplastic
 CC properties and impart gloss to the paper during coating step.
 CC The present sequence represents a GTFB mutant of the invention.
 CC Note: The present sequence is not shown in the specification but
 CC was created by the indexer using the GTFB sequence appearing as AAU98027
 CC and the information in claim 36.
 XX
 XX Sequence 1475 AA;
 SQ
 Query Match 90.0%; Score 99; DB 23; Length 1475;
 Best Local Similarity 90.0%; Pred. No. 1e-07;
 Matches 20; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

DB 552 VPSTFIRAHDSVQDLIRNII 573
 RESULT 11
 AAU98035
 ID AAU98035 standard; Protein; 1475 AA.
 AC AAU98035;
 XX
 XX 27-AUG-2002 (first entry)
 DT
 XX S. mutans glucosyltransferase GTFB mutant D457N/D571K.
 DE
 XX Glucosyltransferase; GTFB; transgenic plant; paper sizing;
 XX coating composition; glucan; starch; latex; thermoplastic molecule;
 XX amyloplast; vacuole; paper manufacture; mutant; mutain.
 KW
 XX Streptococcus mutans.
 OS Synthetic.
 XX
 XX Key Location/Qualifiers
 FH Misc-difference 457
 FT Misc-difference 571
 FT Misc-difference 571
 FT /note= "Wild-type Asp substituted by Asn"
 FT /note= "Wild-type Asp substituted by Lys"
 XX
 XX US2002031826-A1.
 XX
 XX 14-MAR-2002.
 XX
 XX 19-DEC-2000; 2000US-0740274.
 PF
 XX 11-DEC-1998; 98US-0210361.
 XX 07-JUN-1995; 95US-0478704.
 XX 07-JUN-1995; 95US-0482711.
 XX 07-JUN-1995; 95US-0485243.
 XX 16-JAN-1998; 98US-0007999.
 XX 16-JAN-1998; 98US-0008172.
 XX 20-JAN-1998; 98US-0009620.
 XX
 XX (NICH/) NICHOLS S E.
 PA Nichols SE;
 PI WPI; 2002-414332/44.
 DR Glucosyltransferase B or D protein useful for producing a glucan useful
 PT as substitutes for and additions to modified starch and latexes in
 PT paper manufacture, comprises mutations in specific positions -
 XX
 XX Claim 36; Page -; 44pp; English.
 PS
 XX The invention an isolated protein comprising a glucosyltransferase
 CC (GTF) B polypeptide having changes at position from 1448V, D457N,
 CC D567T, K1014T, D457N/D567T, D457N/D571K, D567T/D571K,
 CC D567T/D571K/K1014T, 1448V/D457N/D567T/D571K/K779Q/K1014T,
 CC Y169A/Y170A/Y171A, and K779Q or a GTF D polypeptide having
 CC changes at positions from T589D, T589E, N471D, N471D/T589D, and
 CC N471D/T589E. Also included are a glucan produced by the GTF mutant,
 CC an isolated polynucleotide which encodes P1 or P2, or its complementary
 CC polynucleotide, a ribonucleic acid sequence encoding the GTF mutant,
 CC an expression cassette comprising the polynucleotide operably linked to a
 CC promoter, a vector comprising the expression cassette, host cell
 CC introduced with the vector, a transgenic plant comprising the
 CC vector, a seed or tuber from the transgenic plant, a paper string and/or
 CC coating composition comprising a glucan produced in a plant transformed
 CC with a gene encoding the mutant GTF, wild type or, starch, a latex,
 CC thermoplastic molecule or their combinations or glucan and starch where
 CC the glucan is produced in the amyloplast and/or vacuole or a maize line
 CC deficient in starch biosynthesis, transformed with a gene encoding a
 CC glucosyltransferase B or D enzyme, wild-type or mutant and a paper
 CC comprising the glucan (paper sizing/coating agent). The vector is useful
 CC for producing a glucan in a plant. The method comprises transforming a

CC plant cell with the vector, growing the plant cell under plant growing
 CC conditions to produce a regenerated plant and inducing expression of the
 CC polynucleotide for a time sufficient to produce the glucan in the
 CC regenerated plant, where the vector contains a transit sequence from
 CC ribulose biphosphate carboxylase small subunit, waxy, brittle-1 and
 CC chlorophyll AB binding protein to produce a transgenic plant, and glucan
 CC is produced in the amyloplast of potato or the vacuole of sugar beet.
 CC Glucans are useful as substitutes for and additions to modified starch
 CC and latexes in paper manufacture. Unlike prior art techniques, which
 CC require input materials that produce chemical effluents, paper
 CC manufacture utilizing the glucan produced by GTF, which utilizes
 CC biologically produced input materials, is more cost-effective and
 CC environmentally friendly. Moreover, glucans also exhibit thermoplastic
 CC properties and impart gloss to the paper during coating step.
 CC The present sequence represents a GTFB mutant of the invention.
 CC Note: The present sequence is not shown in the specification but
 CC was created by the indexer using the GTFB sequence appearing as AAU98027
 CC and the information in claim 36.

XX Sequence 1475 AA;

Query Match 89.1%; Score 98; DB 23; Length 1475;
 Best Local Similarity 90.9%; Pred. No. 1.5e-07;
 Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 VPSYSFIRAHDSFVQDLIRNII 22
 |||||
 DB 552 VPSYSFIRAHDSFVQDLIRNII 573

RESULT 12

AA32925
 ID AAR32925 standard; Protein; 1592 AA.

XX AAR32925;

XX 28-JUN-1993 (first entry)

XX Glucosyltransferase I.

XX GT-1; Streptococcus; dental; caries.

XX Streptococcus sobrinus.

XX JP05023188-A.

XX 02-FEB-1993.

XX 25-JUL-1991; 91JP-0186592.

XX 25-JUL-1991; 91JP-0186592.

XX (FUKU/) FUKUI I.

XX (KATO/) KATO K.

XX WPI; 1993-079449/10.

XX N-PDSB; AAQ37760.

XX DNA sequence glucosyltransferase-I - comprises Streptococcus
 PT sobrinus DNA sequence with at least one nucleotide added or
 PT deleted

XX Claim 13; Page 15; 29pp; Japanese.

XX The DNA sequence from Streptococcus sobrinus strain 6715 encodes
 CC glucosyltransferase-I (and mutants). The DNA was obtd. by treating
 CC S. sobrinus 6715 with mutanolysin, extracting the chromosomal DNA,
 CC partially digesting with Sau3AI and fractionating on agarose gel.
 CC The 3-5 kbp fragment was ligated into pUC18 and E. coli JM109
 CC transformed with it. A GT-1 expressing clone was isolated and
 CC sequenced. The clone may be used in the development of a drug for
 CC dental caries.

SQ Sequence 1592 AA;

Query Match 89.1%; Score 98; DB 14; Length 1592;
 Best Local Similarity 86.4%; Pred. No. 1.6e-07;
 Matches 19; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 VPSYSFIRAHDSFVQDLIRNII 22
 |||||
 DB 548 VPSYSFIRAHDSFVQDLIRNII 569

RESULT 13

AAU98032
 ID AAU98032 standard; Protein; 1475 AA.

XX AAU98032;

XX 27-AUG-2002 (first entry)

XX S. mutans glucosyltransferase GTFB mutant D567T.

XX Glucosyltransferase; GTFB; transgenic plant; paper sizing;

XX coating composition; glucan; starch; latex; thermoplastic molecule;

XX amyloplast; vacuole; paper manufacture; mutant; mutain.

XX Streptococcus mutans.

XX Synthetic.

XX Key Location/Qualifiers

XX Misc-difference 567 /note= "Wild-type Asp substituted by Thr"

XX US2002031826-A1.

XX 14-MAR-2002.

XX 19-DEC-2000; 2000US-0740274.

XX 11-DEC-1998; 98US-0210361.

XX 07-JUN-1995; 95US-0478704.

XX 07-JUN-1995; 95US-0482711.

XX 07-JUN-1995; 95US-0485243.

XX 16-JAN-1998; 98US-0007999.

XX 16-JAN-1998; 98US-0008172.

XX 20-JAN-1998; 98US-0009620.

XX (NICH/) NICHOLS S E.

XX Nichols SE;

XX WPI; 2002-414332/44.

XX Claim 36; Page -; 44pp; English.

XX The invention an isolated protein comprising a glucosyltransferase
 CC (GTF) B polypeptide having changes at position 1448V, D457N,
 CC D567T, K1014T, D457N/D567T, D457N/D571K, D567T/D571K,
 CC D567T/D571K/K1014T, I448V/D457N/D567T/D571K/K779Q/K1014T,
 CC Y169N/Y170A/Y171A, and K779Q or a GTF D polypeptide having
 CC changes at positions from T589D, T589E, N471D, N471D/T589D, and
 CC N471D/T589E. Also included are a glucan produced by the GTF mutant,
 CC an isolated polynucleotide which encodes P1 or P2, or its complementary
 CC polynucleotide, a ribonucleic acid sequence encoding the GTF mutant,
 CC an expression cassette comprising the expression cassette, host cell
 CC promoter, a vector comprising a transgenic plant comprising the
 CC introduced with the vector, a transgenic plant comprising the
 CC vector, a seed or tuber from the transgenic plant, a paper sizing and/or
 CC coating composition comprising a glucan produced in a plant transformed
 CC with a gene encoding the mutant GTF, wild type or, starch, a latex,

KM glucosyltransferase; GFPB; transgenic plant; paper sizing;
 KM coating composition; glucan; starch; latex; thermoplastic molecule;
 KM amyloplast; vacuole; paper manufacture; mutant; mutin.
 OS Streptococcus mutans.
 OS Synthetic.
 XX
 FH Key Location/Qualifiers
 FT Misc-difference 567 /note= "Wild-type Asp substituted by Thr"
 FT Misc-difference 571 /note= "Wild-type Asp substituted by Lys"
 FT
 FT US2002031826-A1.
 XX
 XX
 XX 14-MAR-2002.
 PD
 PD 19-DEC-2000; 2000US-0740274.
 XX
 XX 11-DEC-1998; 98US-0210361.
 PR 07-JUN-1995; 95US-0478704.
 PR 07-JUN-1995; 95US-0482711.
 PR 07-JUN-1995; 95US-0485243.
 PR 16-JAN-1998; 98US-0007999.
 PR 16-JAN-1998; 98US-0008172.
 PR 20-JAN-1998; 98US-0009620.
 PR
 XX (NICH/) NICHOLS S E.
 PA
 XX
 XX Nichols SE;
 PI
 DR WPI; 2002-414332/44.
 XX
 PT Glucosyltransferase B or D protein useful for producing a glucan useful
 PT as substitutes for and additions to modified starch and latexes in
 PT paper manufacture, comprises mutations in specific positions -
 XX
 XX Claim 36; Page -; 44pp; English.
 XX
 XX The invention an isolated protein comprising a glucosyltransferase
 CC (GTF) B polypeptide having changes at position from 1448V, D457N,
 CC D567T, K1014T, D457N/D567T, D457N/D571K, D567T/D571K,
 CC D567T/D571K/K1014T, I448V/D457N/D567T/D571K/K799Q/K1014T,
 CC Y169A/Y170A/Y171A, and K779Q or a GTF D polypeptide having
 CC changes at positions from 1589D, T589E, N471D, N471D/T589D, and
 CC N471D/T589E. Also included are a glucan produced by the GTF mutant,
 CC an isolated polynucleotide which encodes P1 or P2, or its complementary
 CC polynucleotide, a ribonucleic acid sequence encoding the GTF mutant,
 CC an expression cassette comprising the polynucleotide operably linked to a
 CC promoter, a vector comprising the expression cassette, host cell
 CC introduced with the vector, a transgenic plant comprising the
 CC vector, a seed or tuber from the transgenic plant, a paper sizing and/or
 CC coating composition comprising a glucan produced in a plant transformed
 CC with a gene encoding the mutant GTF, wild type or, starch, a latex,
 CC thermoplastic molecule or their combinations or glucan and starch where
 CC the glucan is produced in the amyloplast and/or vacuole or a maize line
 CC deficient in starch biosynthesis, transformed with a gene encoding a
 CC glucosyltransferase B or D enzyme, wild-type or mutant and a paper
 CC comprising the glucan (paper sizing/coating agent). The vector is useful
 CC for producing a glucan in a plant. The method comprises transforming a
 CC plant cell with the vector, growing the plant cell under plant growing
 CC conditions to produce a regenerated plant and inducing expression of the
 CC polynucleotide for a time sufficient to produce the glucan in the
 CC regenerated plant, where the vector contains a transit sequence from
 CC ribulose biphosphate carboxylase small subunit, waxy, brittle-1 and
 CC chlorophyll AB binding protein to produce a transgenic plant, and glucan
 CC is produced in the amyloplast of potato or the vacuole of sugar beet.
 CC Glucans are useful as substitutes for and additions to modified starch
 CC and latexes in paper manufacture. Unlike prior art techniques, which
 CC require input materials that produce chemical effluents, paper
 CC manufacture utilizing the glucan produced by GTF, which utilizes
 CC biologically produced input materials, is more cost-effective and
 CC environmentally friendly. Moreover, glucans also exhibit thermoplastic

CC Properties and impact gloss to the paper during coating step.
 CC The present sequence represents a GTFB mutant of the invention.
 CC Note: The present sequence is not shown in the specification but
 CC was created by the indexer using the GTFB sequence appearing as AAU98027
 CC and the information in claim 36.
 CC
 XX
 XX
 SQ Sequence 1475 AA;
 Query Match 82.7%; Score 91; DB 23; Length 1475;
 Best Local Similarity 86.4%; Pred. No. 2.4e-06;
 Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 1 VPESYFIRAHDSFVDLRNII 22
 DB 552 VPESYFIRAHDSFVDTLAKII 573
 Search completed: November 13, 2003, 09:38:28
 Job time : 40.3507 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 13, 2003, 09:31:40 : Search time 13.6588 Seconds
(without alignments)
154.898 Million cell updates/sec

Title: US-09-290-049A-15
Perfect score: 110
Sequence: 1 VPSYSTRADHSEVQDLIRNII 22

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 283308 segs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :
1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	110	100.0	1375	2 JT0345	dextranucrase (EC
2	99	90.0	1475	2 B3135	glfB protein precu
3	98	89.1	1592	2 A38175	glucosyltransferas
4	83	75.5	1290	2 UC5473	dextranucrase (EC
5	83	75.5	1508	2 T31098	probable dextranu
6	76	69.1	1365	2 A41483	glucosyltransferas
7	71	64.5	1449	2 T30857	glucosyltransferas
8	70	63.6	1577	2 T30852	glucosyltransferas
9	68	61.8	1431	2 A45866	dextranucrase (EC
10	68	61.8	1518	2 A44811	glucosyltransferas
11	68	61.8	1599	2 G22737	glucosyltransferas
12	68	61.8	1599	2 G22737	glucosyltransferas
13	50	45.5	51	2 G22737	hypothetical prote
14	48	43.6	540	1 OYHUCR	hypothetical prote
15	46.5	42.3	414	2 T6120	MD2 protein - Yea
16	46	41.8	196	2 S48302	DNA polymerase III
17	46	41.8	597	2 P82935	conserved hypotet
18	46	41.8	631	2 P80599	probable transcrip
19	45.5	41.4	767	2 T39715	conserved hypotet
20	45.5	41.4	956	2 H81654	hypothetical prote
21	45	40.9	136	2 T12823	hypothetical prote
22	45	40.9	1835	2 S46082	urea carboxylase (
23	45	40.9	6359	2 T31679	bactitracin synthe
24	44.5	40.5	954	2 G71496	hypothetical prote
25	44	40.0	202	2 E72688	hypothetical prote
26	44	40.0	282	2 S47795	probable transcrip
27	44	40.0	327	2 H59094	hypothetical prote
28	44	40.0	327	2 D86474	probable RING zinc
29	44	40.0	344	2 JH0511	inositol 2-dehydro

30	44	40.0	570	2 B97160	fibronectin-bindin
31	44	40.0	587	2 S63033	hypothetical prote
32	44	40.0	734	2 A82047	primosomal replica
33	44	40.0	1068	1 A43322	1-phosphatidylinos
34	44	40.0	1068	1 I38110	hypothetical prote
35	44	40.0	1583	2 F97846	nitroreductase fam
36	43	39.1	187	2 H97184	regulatory protease
37	43	39.1	214	1 R6BSCA	hypothetical prote
38	43	39.1	635	2 D86265	hypothetical prote
39	43	39.1	691	2 S48390	hypothetical prote
40	43	39.1	832	2 C86431	hypothetical prote
41	43	39.1	1260	2 A86323	7518.5 protein - A
42	42	38.2	108	2 S75540	protein F14D16.3 (
43	42	38.2	145	2 G69098	hypothetical prote
44	42	38.2	225	2 S19428	SRD1 protein - Yea
45	42	38.2	508	2 T50180	nucleolar protein

ALIGNMENTS

RESULT 1
JT0345
dextranucrase (EC 2.4.1.5) precursor - Streptococcus mutans (strain GS-5)
N:Alternate names: sucrose 6-glucosyltransferase
C:Species: Streptococcus mutans
C:Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 24-Sep-1999
C:Accession: JT0345; C3135
R:Ueda, S.; Shiroza, T.; Kuramitsu, H.K.
Gene 69, 101-109, 1988
A:Title: Sequence analysis of the glfC gene from Streptococcus mutans GS-5.
A:Reference number: JT0345; MUID:89137980; PMID:2976010
A:Accession: JT0345
A:Molecule type: DNA
A:Residues: 1-1375 <UED>
A:Experimental source: GS-5
R:Shiroza, T.; Ueda, S.; Kuramitsu, H.K.
J. Bacteriol. 169, 4263-4270, 1987
A:Title: Sequence analysis of the glfB gene from Streptococcus mutans.
A:Reference number: A33135; MUID:87308013; PMID:3040685
A:Accession: C33135
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-349 <SHI>
A:Cross-references: GB:M17361
C:Genetics:
A:Gene: glfC
C:Function:
A:Description: catalyzes the synthesis of both water-soluble and water-insoluble glucans
A:Superfamily: cpl repeat homology
C:Keywords: duplication; glycosyltransferase; hexosyltransferase
F:1-34/Domain: signal sequence #status predicted <SIG>
F:35-1375/Product: glucosyltransferase #status predicted <MAT>
F:1126-1145/Domain: cpl repeat homology <CP1>
F:1253-1272/Domain: cpl repeat homology <CP2>
F:1318-1337/Domain: cpl repeat homology <CP3>

Query Match 100.0%; Score 110; DB 2; Length 1375;
Best Local Similarity 100.0%; Pred. No. 5.2e-09;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VPSYSTRADHSEVQDLIRNII 22
DB 578 VPSYSTRADHSEVQDLIRNII 599

RESULT 2
B3135
glfB protein precursor - Streptococcus mutans
C:Species: Streptococcus mutans
C:Date: 23-Oct-1990 #sequence_revision 23-Oct-1990 #text_change 15-Oct-1999
C:Accession: B3135; A33128
R:Shiroza, T.; Ueda, S.; Kuramitsu, H.K.

J. Bacteriol. 169, 4263-4270, 1987
 A>Title: Sequence analysis of the glfB gene from Streptococcus mutans.
 A:Reference number: A33135; MUID:87308013; PMID:3040685
 A:Accession: B33135
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-1475 <SH1>
 A:Cross-references: GB:M17361; NID:g153639; PIDN:AAA8586.1; PID:g153640
 R:Shiroza, T.; Ueda, S.; Kuramitsu, H.K.
 Submitted to the Protein Sequence Database, September 1990
 A:Reference number: A33128
 A:Accession: A33128
 A>Status: preliminary; not compared with conceptual translation
 A:Molecule type: DNA
 A:Residues: 1-171,173-641, 'N', 643-1475 <SH2>
 A:Experimental source: strain GS-5
 C:Superfamily: cpl repeat homology
 F:1096-1115/Domain: cpl repeat homology <CP1>
 F:1224-1243/Domain: cpl repeat homology <CP2>
 F:1289-1308/Domain: cpl repeat homology <CP3>
 F:1354-1373/Domain: cpl repeat homology <CP4>
 F:1419-1438/Domain: cpl repeat homology <CP5>

Query Match 90.0%; Score 99; DB 2; Length 1475;
 Best Local Similarity 90.9%; Pred. No. 3,2e-07;
 Matches 20; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 VPSTSFIRAHDSVQDILIRNII 22
 |||||
 DB 552 VPSTSFIRAHDSVQDILIRNII 573

RESULT 3

A38175
 glucosyltransferase precursor - Streptococcus sobrinus

C:Species: Streptococcus sobrinus
 C>Date: 28-Aug-1992 #sequence_revision 28-Aug-1992 #text_change 15-Oct-1999
 C:Accession: A38175
 R:Abou, H.; Mtsunura, T.; Kodama, T.; Ohta, H.; Fukui, K.; Kato, K.; Kagawa, H.
 J. Bacteriol. 173, 989-996, 1991
 A>Title: Peptide sequences for sucrose splitting and glucan binding within Streptococcus
 A:Reference number: A38175; MUID:91123227; PMID:1704006
 A:Accession: A38175
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-1592 <ABO>
 A:Cross-references: GB:D90213; NID:g217032; PIDN:BA14241.1; PID:d1014946; PID:g217033
 C:Superfamily: cpl repeat homology
 F:1093-1112/Domain: cpl repeat homology <CP1>
 F:1222-1241/Domain: cpl repeat homology <CP2>
 F:1287-1306/Domain: cpl repeat homology <CP3>
 F:1330-1351/Domain: cpl repeat homology <CP4>
 F:1352-1371/Domain: cpl repeat homology <CP5>
 F:1402-1420/Domain: cpl repeat homology <CP6>
 F:1465-1484/Domain: cpl repeat homology <CP7>
 F:1513-1532/Domain: cpl repeat homology <CP8>

Query Match 89.1%; Score 98; DB 2; Length 1592;
 Best Local Similarity 86.4%; Pred. No. 5.1e-07;
 Matches 19; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 VPSTSFIRAHDSVQDILIRNII 22
 |||||
 DB 548 VPSTSFIRAHDSVQDILIRNII 569

RESULT 4

JC5473
 dextranucrase (EC 2.4.1.5) - Leuconostoc mesenteroides

C:Species: Leuconostoc mesenteroides
 C>Date: 07-Jul-1997 #sequence_revision 29-Aug-1997 #text_change 29-Aug-1997
 C:Accession: JC5473
 R:Monchois, V.; Willemot, R.M.; Renaud-Simeon, M.; Croux, C.; Monsan, P.

Gene 182, 23-32, 1996
 A>Title: Cloning and sequencing of a gene coding for a novel dextranucrase from Leuconostoc
 A:Reference number: JC5473; MUID:97136686; PMID:8982063
 A:Accession: JC5473
 A>Status: nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-1290 <MON>
 A:Cross-references: GB:U38181
 C:Comment: This enzyme catalyzes the transfer of D-glucopyranosyl units from sucrose onto
 C:Genetics:
 A:Gene: dxtA
 C:Keywords: glycosyltransferase; hexosyltransferase
 F:78-870/Domain: catalytic #status predicted <CAT>
 F:922-1290/Domain: glucan-binding #status predicted <GCB>

Query Match 75.5%; Score 83; DB 2; Length 1290;
 Best Local Similarity 76.2%; Pred. No. 9.8e-05;
 Matches 16; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

OY 2 PNSTSFIRAHDSVQDILIRNII 22
 |||||
 DB 388 PNSTSFIRAHDSVQDILIRNII 408

RESULT 5

T31098
 probable dextranucrase (EC 2.4.1.5), extracellular - Leuconostoc mesenteroides

C:Species: Leuconostoc mesenteroides
 C>Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 11-May-2000
 C:Accession: T31098
 R:Monchois, V.; Renaud-Simeon, M.; Monsan, P.; Willemot, R.M.
 FEBS Microbiol. Lett. 159, 307-315, 1998
 A>Title: Cloning and sequencing of a gene coding for an extracellular dextranucrase (DS
 A:Reference number: Z20981; MUID:98164374; PMID:9503626
 A:Accession: T31098
 A>Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-1508 <MON>
 A:Cross-references: EMBL:AF030129; NID:g2766611; PID:g2766612; PIDN:AA895453.1
 A:Experimental source: strain NRRL B-1299
 C:Genetics:
 A:Gene: dxtB
 C:Function: produces dextran composed only of alpha(1-6) glucosidic bonds
 A:Description: glycosyltransferase; hexosyltransferase
 C:Keywords: glycosyltransferase; hexosyltransferase

Query Match 75.5%; Score 83; DB 2; Length 1508;
 Best Local Similarity 68.2%; Pred. No. 0.00012;
 Matches 15; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

OY 1 PNSTSFIRAHDSVQDILIRNII 22
 |||||
 DB 634 PNSTSFIRAHDSVQDILIRNII 655

RESULT 6

A41483
 glucosyltransferase (EC 2.4.1.-) gtfS precursor - Streptococcus sobrinus

C:Species: Streptococcus sobrinus
 C>Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 15-Oct-1999
 C:Accession: A41483
 R:Gilmore, K.S.; Russell, R.R.B.; Ferretti, J.J.
 Infect. Immun. 58, 2452-2458, 1990
 A>Title: Analysis of the Streptococcus downei gtfS gene, which specifies a glucosyltransferase
 A:Reference number: A41483; MUID:90316665; PMID:2142479
 A:Accession: A41483
 A:Molecule type: DNA
 A:Residues: 1-1365 <GII>
 A:Cross-references: GB:M30943; NID:g153652; PIDN:AAA26898.1; PID:g153653
 C:Genetics:
 A:Gene: gtfS
 C:Superfamily: cpl repeat homology
 C:Keywords: glycosyltransferase; hexosyltransferase

Query Match 69.1%; Score 76; DB 2; Length 1365;
Best Local Similarity 72.7%; Pred. No. 0.0014;
Matches 16; Conservative 1; Mismatches 5; Indels 0; Gaps 0;
QY 1 VPSYFIRAHSEVQDLIRNII 22
DB 537 VNYVFIKRAHSEVQTRIAKII 558

RESULT 7

T30857
glucosyltransferase - Streptococcus salivarius
C:Species: Streptococcus salivarius
C:Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 22-Oct-1999
C:Accession: T30857
R:Simpson, C.L.; Giffard, P.M.; Jacques, N.A.
I: Infect. Immun. 63, 609-621, 1995
A:Title: Streptococcus salivarius ATCC 25975 possesses at least two genes coding for prt
A:Reference number: Z20909; MUID:95122197; PMID:7822030
A:Accession: T30857
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Residues: 1-1449 <SIM>
A:Molecule type: DNA
A:Cross-references: EMBL:L35495; NID:g662378; PID:g662379; PIDN:AA041412.1
C:Genetics:
A:Gene: gtfN

Query Match 64.5%; Score 71; DB 2; Length 1449;
Best Local Similarity 65.0%; Pred. No. 0.0093;
Matches 13; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 3 SYSFIRAHSEVQDLIRNII 22
DB 609 NYAFVRAHSEVQSIIGQII 628

RESULT 8

T30552
glucosyltransferase N - Streptococcus salivarius (fragment)
C:Species: Streptococcus salivarius
C:Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 22-Oct-1999
C:Accession: T30552
R:Jaffe, R.I.
I: submitted to the EMBL Data Library, February 1998
A:Description: Streptococcus salivarius V1477 gtfN.
A:Reference number: Z20854
A:Accession: T30552
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1449 <JAF>
A:Cross-references: EMBL:AF049609; NID:g2935545; PID:g2935546; PIDN:AA005156.1
C:Genetics:
A:Gene: gtfN

Query Match 64.5%; Score 71; DB 2; Length 1449;
Best Local Similarity 65.0%; Pred. No. 0.0093;
Matches 13; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 3 SYSFIRAHSEVQDLIRNII 22
DB 609 NYAFVRAHSEVQSIIGQII 628

RESULT 9

T30858
glucosyltransferase - Streptococcus salivarius
C:Species: Streptococcus salivarius
C:Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 22-Oct-1999
C:Accession: T30858
R:Simpson, C.L.; Giffard, P.M.; Jacques, N.A.
I: Infect. Immun. 63, 609-621, 1995
A:Title: Streptococcus salivarius ATCC 25975 possesses at least two genes coding for prt

A:Reference number: Z20909; MUID:95122197; PMID:7822030
A:Accession: T30858
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1577 <SIM>
A:Cross-references: EMBL:L35928; NID:g662380; PID:g662381; PIDN:AA041413.1
C:Genetics:
A:Gene: gtfN

Query Match 63.6%; Score 70; DB 2; Length 1577;
Best Local Similarity 65.0%; Pred. No. 0.015;
Matches 13; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 3 SYSFIRAHSEVQDLIRNII 22
DB 661 NYFVRAHSEVQAVLANII 680

RESULT 10

A45866
dextranucrase (EC 2.4.1.5) precursor - Streptococcus mutans
C:Species: Streptococcus mutans
C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 07-May-1999
C:Accession: A45866
R:Honda, O.; Kato, C.; Kuramitsu, H.K.
U: Gen. Microbiol. 136, 2099-2105, 1990
A:Title: Nucleotide sequence of the Streptococcus mutans gtfD gene encoding the glucosyl
A:Reference number: A45866; MUID:91100958; PMID:2148600
A:Accession: A45866
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1431 <HON>

A:Cross-references: GB:M29296
C:Superfamily: cpl repeat homology
C:Keywords: glycosyltransferase; hexosyltransferase
F:181-201/Domain: cpl repeat homology <CP1>
F:1127-1146/Domain: cpl repeat homology <CP2>
F:1192-1211/Domain: cpl repeat homology <CP3>
F:1257-1276/Domain: cpl repeat homology <CP4>
F:1277-1297/Domain: cpl repeat homology <CP5>
F:1321-1340/Domain: cpl repeat homology <CP6>
F:1341-1361/Domain: cpl repeat homology <CP6>
F:1385-1404/Domain: cpl repeat homology <CP7>

Query Match 61.8%; Score 68; DB 2; Length 1431;
Best Local Similarity 70.0%; Pred. No. 0.028;
Matches 14; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 3 SYSFIRAHSEVQDLIRNII 22
DB 576 NYFIRAHSEVQTVIAKII 595

RESULT 11

A44811
glucosyltransferase (EC 2.4.1.-) I - Streptococcus salivarius
C:Species: Streptococcus salivarius
C:Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 15-Oct-1999
C:Accession: A44811; S22726; S28809
R:Giffard, P.M.; Simpson, C.L.; Milward, C.P.; Jacques, N.A.
U: Gen. Microbiol. 137, 2577-2593, 1991
A:Title: Molecular characterization of a cluster of at least two glucosyltransferase gen
A:Reference number: A44811; MUID:92148377; PMID:1838391
A:Accession: A44811
A:Molecule type: DNA
A:Residues: 1-1518 <GIF>
A:Cross-references: EMBL:Z11873; NID:g47526; PIDN:CAA77900.1; PID:g47527
A:Note: sequence extracted from NCBI backbone (NCBIN:81050, NCBIP:81052)
C:Genetics:
A:Gene: gtfU
C:Superfamily: cpl repeat homology
C:Keywords: glycosyltransferase; hexosyltransferase
F:1307-1326/Domain: cpl repeat homology <CP4>

Query Match 61.8%; Score 68; DB 2; Length 1518;
Best Local Similarity 65.0%; Pred. No. 0.03;
Matches 13; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 3 SYSFIRAHSEVODLIRNII 22
DB 604 NYVFRADHNNQDIIAEII 623

RESULT 12
S22737
glucosyltransferase (EC 2.4.1.-) S - Streptococcus salivarius

C/Species: Streptococcus salivarius
C/Date: 16-Sep-1992 #sequence_revision 16-Sep-1992 #text_change 21-Jan-2000
C/Accession: S22737; S28810; B44811; S22727
R/Jacques, N.

Submitted to the EMBL Data Library, March 1992
A/Reference number: S22726

A/Accession: S22737

A/Molecule type: DNA
A/Residues: 1-1599 <JAC>

A/Cross-references: EMBL:Z11872; NID:G47530; PIDN:CAA7898.1; PID:G47531

A/Experimental source: ATCC 25975
R/Giffard, P.M.; Simpson, C.L.; Milward, C.P.; Jacques, N.A.

J. Gen. Microbiol. 137, 2577-2593, 1991
A/Title: Molecular characterization of a cluster of at least two glucosyltransferase genes

A/Reference number: A44811; MUID:Z2148377; PMID:1838331

A/Accession: S28810
A/Molecule type: DNA
A/Residues: 1-51 <GIF>

A/Cross-references: EMBL:Z11873
C/Genetics:

A/Gene: gtfK
C/Superfamily: cpl repeat homology
C/Keywords: glycosyltransferase; hexosyltransferase

F/1456-1475/Domain: cpl repeat homology <CPR>

Query Match 61.8%; Score 68; DB 2; Length 1599;
Best Local Similarity 65.0%; Pred. No. 0.031;
Matches 13; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 3 SYSFIRAHSEVODLIRNII 22
DB 574 TYLFRADHSEVQVIADII 593

RESULT 13
G82455
hypothetical protein VCA0471 [imported] - Vibrio cholerae (strain N16961 serogroup O1)

C/Species: Vibrio cholerae
C/Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Feb-2001

C/Accession: G82455
R/Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwin, M.L.; Dodeson, R.J.;

Charadon, D.; Ermolaeva, M.D.; Yamahavean, J.; Bass, S.; Qin, H.; Drogot, I.; Sellers, F.
1, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.

Nature 406, 477-483, 2000
A/Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.

A/Reference number: A82035; MUID:20406853; PMID:10952301

A/Accession: G82455

A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-51 <HRI>

A/Cross-references: GB:AE004379; GB:AE003853; NID:G9657865; PIDN:AAF96375.1; GSPDB:GN001
A/Experimental source: serogroup O1; strain N16961; biotype El Tor

C/Genetics:
A/Gene: VCA0471
A/Map position: 2

Query Match 45.5%; Score 50; DB 2; Length 51;
Best Local Similarity 54.5%; Pred. No. 0.43;
Matches 12; Conservative 5; Mismatches 3; Indels 2; Gaps 1;

QY 1 VPSYFIRAHSEVODLIRNII 22
DB 15 VPAVSALR--NSHFRDLKEXEI 34

RESULT 14
OYHUCR
natriuretic peptide receptor C precursor - human

N/Alternate names: atrial natriuretic factor clearance receptor; natriuretic peptide cle
C/Species: Homo sapiens (man)
C/Date: 09-Nov-1990 #sequence_revision 05-May-1995 #text_change 22-Jun-1999

C/Accession: S10150; A35896
R/Lowe, D.G.; Camerato, T.R.; Goeddel, D.V.

Nucleic Acids Res. 18, 3412, 1990
A/Title: cDNA sequence of the human atrial natriuretic peptide clearance receptor.

A/Reference number: S10150; MUID:90287735; PMID:2162522

A/Accession: S10150
A/Molecule type: mRNA

A/Residues: 1-540 <LOW>
A/Cross-references: EMBL:X52282; NID:G28705; PIDN:CAA36523.1; PID:G28706

A/Note: alternative splice form C6
R/Porter, J.G.; Arfsten, A.; Fuller, F.; Miller, J.A.; Gregory, L.C.; Lewicki, J.A.

Biochem. Biophys. Res. Commun. 171, 796-803, 1990
A/Title: Isolation and functional expression of the human atrial natriuretic peptide cle

A/Reference number: A35896; MUID:90386656; PMID:2169733

A/Accession: A35896
A/Molecule type: mRNA

A/Residues: 1-475, 'SG', 477-540 <POR>
A/Cross-references: GB:M59305; NID:G178651; PIDN:AAA51734.1; PID:G178652

A/Note: alternative splice form C5
R/Shultz, J.T.; O'Connell, K.L.; Garcia, C.; Wong, S.; Engel, A.M.; Garbers, D.L.; Lowe,

Biochemistry 33, 11372-11381, 1994
A/Title: The disulfide linkages and glycosylation sites of the human natriuretic peptide

A/Reference number: A55870; MUID:95244450; PMID:7727388

A/Contents: annotation; propeptide, disulfide bonds, carbohydrate binding sites
C/Genetics:

A/Gene: GDB:NR3; NPRC; ANPRC
A/Cross-references: GDB:125201; OMIM:108962

A/Map position: 5p14-5p13
C/Superfamily: natriuretic peptide receptor C; natriuretic peptide-binding domain homolo

C/Keywords: alternative splicing; glycoprotein; homodimer; hormone receptor; transmembr

F/1-23/Domain: signal sequence #status predicted <SIG>
F/24-45/Domain: propeptide #status experimental <PRO>

F/46-540/Product: natriuretic peptide receptor C #status predicted <EXT>
F/46-479/Domain: extracellular #status predicted <EXT>

F/481-503/Domain: transmembrane #status predicted <TM>
F/131-462/Domain: natriuretic peptide-binding domain homology <NDB>

F/86,293/Binding site: carbohydrate (asn) (covalent) #status experimental
F/108-136,213-261/Disulfide bonds: #status experimental (partial) #status experimental

F/394/Binding site: carbohydrate (asn) (covalent) (in splice form C6) #status experimental
F/473-476/Disulfide bonds: (partial) (in splice form C6) #status experimental

Query Match 43.6%; Score 48; DB 1; Length 540;
Best Local Similarity 44.4%; Pred. No. 14;
Matches 8; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

QY 4 YSFIRAHSEVODLIRNII 21
DB 234 YSFDETKDLDIEDIYRNI 251

RESULT 15
T06120
hypothetical protein F23E12.70 - Arabidopsis thaliana

C/Species: Arabidopsis thaliana (mouse-ear cress)
C/Date: 30-Apr-1999 #sequence_revision 30-Apr-1999 #text_change 22-Oct-1999

C/Accession: T06120
R/Bevan, M.; Hilbert, H.; Braun, M.; Holzer, B.; Brandt, A.; Duesterhoeft, A.; Hobeisel,

submitted to the Protein Sequence Database, April 1999
A/Reference number: Z15485

A/Accession: T06120

A/Molecule type: DNA
A/Residues: 1-414 <BEV>

A:Cross-references: EMBL:AL022604; GSPDB:GN00062; ATSP:F23E12.70
 A:Experimental source: cultivar Columbia; BAC clone F23E12
 C:Genetics:
 A:Gene: ATSP:F23E12.70
 A:Map position: 4
 A:introns: 83/1; 109/3; 180/1; 199/3; 215/3; 259/3; 284/3; 317/3; 359/3
 Query Match 42.3%; Score 46.5; DB 2; Length 414;
 Best Local Similarity 55.0%; Pred. No. 18;
 Matches 11; Conservative 2; Mismatches 6; Indels 1; Gaps 1;
 QY 2 PSYFIRAHDSRVQDIIRNI 21
 ||: |||||:
 Db 337 PSF-IIHAHSEVSSISYNI 355
 Search completed: November 13, 2003, 09:50:28
 Job time : 13.6588 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 13, 2003, 09:31:40 / Search time 7.92417 Seconds
(without alignments)
130.561 Million cell updates/sec

Title: US-09-290-049a-15

Perfect score: 110
Sequence: 1 VPSYSFIRAHSEVQDLIRNII 22

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	110	100.0	1455	1	GFPC_STRMU
2	105	95.5	1476	1	GFPC_STRMU
3	100	90.9	1597	1	GFPL_STRDO
4	98	89.1	1592	1	GFPL_STRDO
5	96	89.1	1365	1	GFPL_STRDO
6	68	61.8	1462	1	GFPL_STRDO
7	48	43.6	541	1	ANPC_HUMAN
8	46	41.8	1835	1	DURL_YEAST
9	45	40.9	6359	1	BACC_BACCI
10	45	40.9	282	1	YTAU_ECOLI
11	44	40.0	344	1	YTAU_ECOLI
12	44	40.0	1068	1	YTAU_ECOLI
13	44	40.0	1068	1	YTAU_ECOLI
14	44	40.0	1068	1	YTAU_ECOLI
15	44	40.0	1068	1	YTAU_ECOLI
16	44	40.0	1068	1	YTAU_ECOLI
17	44	40.0	1068	1	YTAU_ECOLI
18	44	40.0	1068	1	YTAU_ECOLI
19	44	40.0	1068	1	YTAU_ECOLI
20	44	40.0	1068	1	YTAU_ECOLI
21	44	40.0	1068	1	YTAU_ECOLI
22	44	40.0	1068	1	YTAU_ECOLI
23	44	40.0	1068	1	YTAU_ECOLI
24	44	40.0	1068	1	YTAU_ECOLI
25	44	40.0	1068	1	YTAU_ECOLI
26	44	40.0	1068	1	YTAU_ECOLI
27	44	40.0	1068	1	YTAU_ECOLI
28	44	40.0	1068	1	YTAU_ECOLI
29	44	40.0	1068	1	YTAU_ECOLI
30	44	40.0	1068	1	YTAU_ECOLI
31	44	40.0	1068	1	YTAU_ECOLI
32	44	40.0	1068	1	YTAU_ECOLI
33	44	40.0	1068	1	YTAU_ECOLI

34	41	37.3	659	1	RNB_HAEN
35	41	37.3	826	1	TLR4_PAPAN
36	41	37.3	875	1	ENV_BIV06
37	41	37.3	904	1	ENV_BIV27
38	41	37.3	1006	1	K125_TOBAC
39	41	37.3	2542	1	TLN2_HUMAN
40	41	37.3	63	1	VTX_BYDVP
41	40.5	36.8	183	1	RET_BOVIN
42	40.5	36.8	201	1	RET_PIG
43	40.5	36.8	96	1	IBBA_PEA
44	40.5	36.8	114	1	IBB2_PEA
45	40.5	36.8	205	1	ADEN_ADE68

ALIGNMENTS

RESULT 1
GFPC_STRMU STANDARD; PRT: 1455 AA.
ID GFPC_STRMU
AC P13470; 069382; 069385; 069388; 069391; 069397; P05427;
DR 01-NOV-1988 (Rel. 09, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DI 15-SEP-2003 (Rel. 42, Last annotation update)
DE Glucosyltransferase-SI precursor (EC 2.4.1.5) (GTF-SI)
DE (Dextranucrase) (Sucrose 6-glucosyltransferase).
GN GFPC OR SWU.1005.
OS Streptococcus mutans.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1309;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=GS-5;
RX MEDLINE=89137980; PubMed=2976010;
RA Ueda S., Shiroza T., Kuramitsu H.K.;
RT "Sequence analysis of the gfpc gene from Streptococcus mutans GS-5."
RL Gene 69:101-109(1988).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=MT4239 / Serotype C, MT4245 / Serotype E, MT4251 / Serotype F,
RX MT4467 / Serotype B, and MT4148 / Serotype C;
RA MEDLINE=98231643; PubMed=9570124;
RA Fujiwara T., Terao Y., Hoshino T., Kawabata S., Oshima T., Sobue S.,
RA Kimura S., Hamada S.;
RT "Molecular analyses of glucosyltransferase genes among strains of
RT Streptococcus mutans."
RL FEWS Microbiol. Lett. 161:331-336(1998).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=U0159 / ATCC 700610 / Serotype C;
RX MEDLINE=22295063; PubMed=12397186;
RA Ajdic D., McShan W.M., McLaughlin R.E., Savic G., Chang J., Qian X.,
RA Carson M.B., Primeaux C., Tian R., Kenton S., Dia H., Lin S., Qian X.,
RA Li S., Zhu H., Najjar F., Lai H., White J., Roe B.A., Perleth J.J.;
RT "Genome sequence of Streptococcus mutans U0159, a cariogenic dental
RT pathogen."
RL Proc. Natl. Acad. Sci. U.S.A. 99:14434-14439(2002).
RN [4]
RP SEQUENCE OF 1-349 FROM N.A.
RC STRAIN=GS-5;
RX MEDLINE=87308013; PubMed=3040685;
RA Shiroza T., Ueda S., Kuramitsu H.K.;
RT "Sequence analysis of the gfpc gene from Streptococcus mutans."
RN [5]
RP FUNCTION: PRODUCTION OF EXTRACELLULAR GLUCANS, THAT ARE THOUGHT
TO PLAY A KEY ROLE IN THE DEVELOPMENT OF THE DENTAL PLAQUE BECAUSE
OF THEIR ABILITY TO ADHERE TO SMOOTH SURFACES AND MEDIANE THE
AGGREGATION OF BACTERIAL CELLS AND FOOD DEBRIS.
CC AGGREGATION OF BACTERIAL CELLS AND FOOD DEBRIS.
CC -1- CARBOLYTIC ACTIVITY: Sucrose + ((1,6)-alpha-D-glucosyl-)) (N) = D-
CC fructose + ((1,6)-alpha-D-glucosyl-)) (N+1).
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- DISEASE: DENTAL CARIES.

CC -1- MISCELLANEOUS: GTF-1 SYNTHESIZES WATER-INSOLUBLE GLUCANS (ALPHA
CC 1,3-LINKED GLUCOSE AND SOME 1,6 LINKAGES), GTF-S SYNTHESIZES
CC WATER-SOLUBLE GLUCANS (ALPHA 1,6-GLUCOSE), GTF-SI SYNTHESIZES BOTH
CC FORMS OF GLUCANS.
CC -1- SIMILARITY: TO OTHER GLUCOSYLTRANSFERASES AND SOME TO A GLUCAN-
CC BINDING PROTEIN FROM S. MUTANS.
CC -1- SIMILARITY: Contains 5 cell wall binding repeats.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.ebi.ac.uk/announcements>
CC or send an email to license@ebi.ac.uk).

CC -----
CC EMBL; M22054; AAA89592.1; -
CC EMBL; D88652; BAA26102.1; -
CC EMBL; D88655; BAA26106.1; -
CC EMBL; D88658; BAA26110.1; -
CC EMBL; D88661; BAA26114.1; -
CC EMBL; D88678; BAA26120.1; -
CC EMBL; AE014940; AAN58706.1; -
CC EMBL; M17361; AAA86589.1; -
CC InterPro; IPR002479; CW-binding.
CC InterPro; IPR003318; Glyco_hydro_70.
CC Pfam; PF01473; CW-binding_1; 1.
CC Pfam; PF02324; Glyco_hydro_70; 1.
CC Transferrase; Glycosyltransferase; Signal; Repeat; Dental caries;
CC Complete proteome.
CC KW
CC FT SIGNAL 1 34
CC FT CHAIN 35 1455
CC FT DOMAIN 35 1050
CC FT DOMAIN 1126 1455
CC FT DOMAIN 1126 1455
CC FT REPEAT 1126 1159
CC FT REPEAT 1169 1200
CC FT REPEAT 1227 1238
CC FT REPEAT 1253 1303
CC FT REPEAT 1318 1330
CC FT REPEAT 21 21
CC FT VARIANT 81 81
CC FT VARIANT 106 106
CC FT VARIANT 116 116
CC FT VARIANT 126 126
CC FT VARIANT 150 151
CC FT VARIANT 256 256
CC FT VARIANT 425 425
CC FT VARIANT 519 519
CC FT VARIANT 538 538
CC FT VARIANT 545 545
CC FT VARIANT 597 597
CC FT VARIANT 600 600
CC FT VARIANT 601 601
CC FT VARIANT 614 614
CC FT VARIANT 727 727
CC FT VARIANT 734 734
CC FT VARIANT 964 964
CC FT VARIANT 1113 1113
CC FT VARIANT 1118 1118
CC FT VARIANT 1204 1204
CC FT VARIANT 1208 1208
CC FT VARIANT 1294 1294
CC FT VARIANT 1305 1305
CC FT VARIANT 1326 1326
CC FT VARIANT 1331 1331
CC FT VARIANT 1377 1377

GLUCOSYLTRANSFERASE-ST.
CATALYTIC (APPROXIMATE).
GLUCAN-BINDING (APPROXIMATE).
2-4 A, 1 C AND 1 AC REPEATS.
A REPEAT.
A REPEAT.
C REPEAT.
AC REPEAT.
A REPEAT (INCOMPLETE).
V -> I (IN STRAIN GS-5).
P -> L (IN STRAIN MT4239).
D -> V (IN STRAIN GS-5).
S -> A (IN STRAINS GS-5 AND MT4467).
A -> T (IN STRAIN GS-5).
SR -> PK (IN STRAINS GS-5, MT4239 AND
MT4467).
A -> V (IN STRAINS GS-5 AND MT4467).
R -> N (IN STRAIN MT4251).
Y -> D (IN STRAINS MT4245 AND MT4251).
R -> K (IN STRAINS MT4245 AND MT4251).
Y -> F (IN STRAINS MT4245 AND MT4251).
N -> D (IN STRAINS MT4245, MT4251, MT4467
AND MT8148).
R -> K (IN STRAINS MT4245, MT4251, MT4467
AND MT8148).
A -> T (IN STRAIN GS-5).
M -> T (IN STRAIN GS-5).
T -> I (IN STRAIN MT8148).
A -> V (IN STRAIN MT8148).
L -> F (IN STRAIN MT4239).
N -> Y (IN STRAIN MT4239).
A -> T (IN STRAIN MT4239).
I -> V (IN STRAINS GS-5, MT4239, MT4467
AND MT8148).
V -> I (IN STRAIN MT8148).
DOH -> NGY (IN STRAINS GS-5, MT4467 AND
MT8148).
MISSING (IN STRAIN MT4245).
I -> V (IN STRAIN MT8148).
T -> A (IN STRAINS GS-5, MT4239, MT4467
AND MT8148).
R -> K (IN STRAIN MT8148).

FT VARIANT 1398 1398 V -> I (IN STRAIN MT8148).
FT VARIANT 1424 1424 D -> N (IN STRAIN MT4239).
FT VARIANT 1429 1429 V -> I (IN STRAIN MT4239 AND MT8148).
FT VARIANT 1444 1444 S -> P (IN STRAIN MT8148).
FT CONFLICT 1337 1455 ORLYKSNQVQAKGELITKRIKTYDPNSQNEVRYR
TSSGMWYFGNDGVALIGHVVGGRVVFDENGVRASHD
ORHNDYDRDPFGSSAVRRFRSRNCFDNRFF ->
HASTILMYERLRRESLQGVKVSNTMILIPMKFVIVM
(IN REF. 1).
FT FT
FT FT
SQ SEQUENCE 1455 AA; 162965 MW; 3CB455A9A4FE06 CRC64;
Query Match 100.0%; Score 110; DB 1; Length 1455;
Best Local Similarity 100.0%; Pred. NO. 2.6e-09;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 VPYSFIRHDSVQDLIRNII 22
Db 578 VPYSFIRHDSVQDLIRNII 599

RESULT 2
GTFB STRMU STANDARD; PRT; 1476 AA.
ID GTFB STRMU 069381; 069384; 069387; 069390; 069396;
AC P08987; 069381 (Rel. 09, Created)
DT 01-NOV-1988 (Rel. 09, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Glucosyltransferase-I precursor (EC 2.4.1.5) (GTF-I) (Dextranucrase)
DE (Sucrose 6-glucosyltransferase).
DE GTFB OR SMU.1004.
GN Streptococcus mutans.
OS Streptococcus mutans.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1309;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-GS-5;
RX MEDLINE=87308013; PubMed=3040685;
RA Shiroza T., Ueda S., Kuramitsu H.K.;
RT "Sequence analysis of the gtfb gene from Streptococcus mutans."
RL J. Bacteriol. 169:4263-4270(1987).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-MT4239 / Serotype C, MT4245 / Serotype E, MT4251 / Serotype F,
MT4467 / Serotype E, and MT8148 / Serotype C;
RX MEDLINE=98231643; PubMed=9570124;
RA Fujiwara T., Terao Y., Hoshino T., Kawabata S., Ooshima T., Sobue S.,
RA Kimura S., Hamada S.;
RT "Molecular analyses of glucosyltransferase genes among strains of
Streptococcus mutans."
RL FEMS Microbiol. Lett. 161:331-336(1998).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=UA159 / ATCC 700610 / Serotype C;
RX MEDLINE=22295063; PubMed=12397186;
RA Ajdic D., Moshan W.M., McLaughlin R.E., Savic G., Chang J.,
RA Carson M.B., Primeaux C., Tian R., Kenton S., Jia H., Lin S., Qian Y.,
RA Li S., Zhu H., Nejer J., Lai H., White J., Roe B.A., Ferretti J.D.;
RT "Genome sequence of Streptococcus mutans UA159, a cariogenic dental
pathogen."
RL Proc. Natl. Acad. Sci. U.S.A. 99:14434-14439(2002).
CC -1- FUNCTION: PRODUCTION OF EXTRACELLULAR GLUCANS, THAT ARE THOUGHT
CC TO PLAY A KEY ROLE IN THE DEVELOPMENT OF THE DENTAL PLAQUE BECAUSE
CC OF THEIR ABILITY TO ADHERE TO SMOOTH SURFACES AND MEDIANE THE
CC AGGREGATION OF BACTERIAL CELLS AND FOOD DEBRIS.
CC -1- CATALYTIC ACTIVITY: Sucrose + ((1,6)-alpha-D-glucosyl)((N) = D-
CC fructose + ((1,6)-alpha-D-glucosyl)((N+1)).
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- DISEASE: DENTAL CARIES.
CC -1- MISCELLANEOUS: GTF-1 SYNTHESIZES WATER-INSOLUBLE GLUCANS (ALPHA
CC 1,3-LINKED GLUCOSE AND SOME 1,6 LINKAGES), GTF-S SYNTHESIZES
CC WATER-SOLUBLE GLUCANS (ALPHA 1,6-GLUCOSE), GTF-SI SYNTHESIZES BOTH
CC FORMS OF GLUCANS.

FT	VARIANT	1394	1394	Y -> H (IN STRAINS GS-5, MT4239, MT4467 AND MT9148).
FT	VARIANT	1402	1402	S -> G (IN STRAINS GS-5, MT4239, MT4467 AND MT8148).
FT	VARIANT	1459	1459	Y -> H (IN STRAIN MT4467).
FT	CONFLICT	570	570	R -> A (IN REF. 1).
FT	CONFLICT	800	817	ADQDVRAVASTPSTDGK -> LIRKPALRLARPHQQMA (IN REF. 1).
FT	CONFLICT	1310	1310	H -> L (IN REF. 1).
SO	SEQUENCE	1476 AA;	165846 MM;	9C6E09F731B4CCFC CRC64;
Query Match		95.5%;	Score 105;	DB 1; Length 1476;
Best Local Similarity		95.5%;	Pred. No. 1.6e-08;	
Matches	21; Conservative	1;	Mismatches	0; Indels
			Gaps	0;
Dy	1 VPSSYFRHDSSEVDLIRNII 22			
Db	552 VPSYFRHDSSEVDLIRDII 573			
RESULT 3				
GFPI_STRDO	STANDARD;	PRT;	1597 AA.	
ID_GFP1_SPRO				
AC	P11001			
DT	01-VUL-1989 (Rel. 11, Created)			
DT	01-VUL-1989 (Rel. 11, Last sequence update)			
DT	15-SEP-2003 (Rel. 42, Last annotation update)			
DE	Glucosyltransferase-I precursor (EC 2.4.1.5) (GTF-I) (Dextranucrase)			
DE	(Sucrose 6-glucosyltransferase).			
GN	GTFI.			
OS	Streptococcus downei (Streptococcus sobrinus).			
OC	Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;			
OX	NCHI_TaxID=1317;			
RP	[1]			
RN	SEQUENCE FROM N.A.			
RC	STRAIN=MFE28;			
RX	MEDLINE=87308014; PubMed=3040686;			
RA	Ferrelli J.J., Gilpin M.L., Russell R.R.B.;			
RT	"Nucleotide sequence of a glucosyltransferase gene from Streptococcus sobrinus MFE28";			
RL	J. Bacteriol. 169:4271-4278 (1987).			
CC	- FUNCTION: PRODUCTION OF EXTRACELLULAR GLUCANS, THAT ARE THOUGHT TO PLAY A KEY ROLE IN THE DEVELOPMENT OF THE DENTAL PLAQUE BECAUSE OF THEIR ABILITY TO ADHERE TO SMOOTH SURFACES AND MEDIANE THE AGGREGATION OF BACTERIAL CELLS AND FOOD DEBRIS.			
CC	- CATALYTIC ACTIVITY: Sucrose + {(1,6)-alpha-D-glucosyl}(N) = D-fructose + {(1,6)-alpha-D-glucosyl}(N+1).			
CC	- SUBCELLULAR LOCATION: secreted.			
CC	- DISEASE: DENTAL CARIES.			
CC	- MISCELLANEOUS: GTF-I SYNTHESIZES WATER-INSOLUBLE GLUCANS (ALPHA 1,3-LINKED GLUCOSE AND SOME 1,6 LINKAGES), GTF-S SYNTHESIZES WATER-SOLUBLE GLUCANS (ALPHA 1,6-GLUCOSE). GTF-SI SYNTHESIZES BOTH FORMS OF GLUCANS.			
CC	- SIMILARITY: TO OTHER GLUCOSYLTRANSFERASES AND SOME TO A GLUCAN-BINDING PROTEIN FROM S.MUTANS.			
CC	- SIMILARITY: Contains 19 cell wall binding repeats.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).			
CC	-----			
DR	EMBL; M17391; AAC63063.1; -			
DR	InterPro; IPR002479; CM_binding.			
DR	InterPro; IPR003318; Glyco_hydro_70.			
DR	Pfam; PF01473; CM_binding_1; 16.			
DR	Pfam; PF03324; Glyco_hydro_70; 1.			
KW	Transferase; Glucosyltransferase; Signal; Repeat; Dental caries.			
FT	SIGNAL	1	38	POTENTIAL.

```

FT CHAIN 39 1597 GLUCOSYLTRANSFERASE-1.
FT DOMAIN 39 1050 CATALYTIC (APPROXIMATE).
FT DOMAIN 1099 1597 GLUCAN-BINDING (APPROXIMATE).
FT DOMAIN 1099 1597 TRANSFERASE; Glycosyltransferase; Signal; Repeat; Dental carries.
FT REPEAT 1099 1597 1.25 A, 2 B, AND 5 AC REPEATS.
FT REPEAT 1163 1233 A REPEAT.
FT REPEAT 1163 1233 AC REPEAT.
FT REPEAT 1227 1277 AC REPEAT.
FT REPEAT 1227 1277 AC REPEAT.
FT REPEAT 1292 1342 B REPEAT.
FT REPEAT 1292 1342 AC REPEAT.
FT REPEAT 1352 1399 B REPEAT.
FT REPEAT 1352 1399 AC REPEAT.
FT REPEAT 1406 1455 B REPEAT.
FT REPEAT 1406 1455 AC REPEAT.
FT REPEAT 1465 1512 B REPEAT.
FT REPEAT 1465 1512 AC REPEAT.
FT REPEAT 1519 1568 B REPEAT.
FT REPEAT 1519 1568 AC REPEAT.
FT REPEAT 1582 1597 A REPEAT (INCOMPLETE).
SQ SEQUENCE 1597 AA; B9B6A200868798E CRC64;

Query Match 90.9%; Score 100; DB 1; Length 1597;
Best Local Similarity 90.9%; Pred. No. 1,1e-07;
Matches 20; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 VPSYSPFRAHDSVODLRNII 22
Db 554 VPSYSPFRAHDSVODLRDII 575

RESULT 4
GTF2_STRDO STANDARD; PRT; 1592 AA.
ID GTF2_STRDO
AC P27470;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Glucosyltransferase-I precursor (EC 2.4.1.5) (GTF-I) (Dextranucrase)
DE (Sucrose 6-glucosyltransferase).
OS Streptococcus downei (Streptococcus sobrinus).
DE Streptococcus downei (Streptococcus sobrinus);
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OC NCBI_TaxID=1317;
OK NCBI_TaxID=1317;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=6715 / Serotype G;
RX MEDLINE=91123227; PubMed=1704006;
RA Abo H., Matsumura T., Kodama T., Ohta H., Fukui K., Kato K.,
RA Kagawa H.;
RT "Peptide sequences for sucrose splitting and glucan binding within
RT Streptococcus sobrinus glucosyltransferase (water-insoluble glucan
RT synthetase).";
RL J. Bacteriol. 173:989-996 (1991).
CC -1- FUNCTION: PRODUCTION OF EXTRACELLULAR GLUCANS, THAT ARE THOUGHT
CC TO PLAY A KEY ROLE IN THE DEVELOPMENT OF THE DENTAL PLAQUE BECAUSE
CC OF THEIR ABILITY TO ADHERE TO SMOOTH SURFACES AND MEDIANE THE
CC AGGREGATION OF BACTERIAL CELLS AND FOOD DEBRIS.
CC -1- CATALYTIC ACTIVITY: Sucrose + ((1,6)-alpha-D-glucosyl)(N) = D-
CC fructose + ((1,6)-alpha-D-glucosyl)(N+1).
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- DISEASE: DENTAL CARRIES.
CC -1- MISCELLANEOUS: GTF-I SYNTHESIZES WATER-INSOLUBLE GLUCANS (ALPHA
CC 1,3-LINKED GLUCOSE AND SOME 1,6 LINKAGES), GTF-S SYNTHESIZES BOTH
CC WATER-SOLUBLE GLUCANS (ALPHA 1,6-GLUCOSE). GTF-SI SYNTHESIZES BOTH
CC FORMS OF GLUCANS.
CC -1- SIMILARITY: TO OTHER GLUCOSYLTRANSFERASES AND SOME TO A GLUCAN-
CC BINDING PROTEIN FROM S.MUTANS.
CC -1- SIMILARITY: Contains 16 cell wall binding repeats.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; D90213; BA14241.1;
CC InterPro; IPR002479; CW_binding.

```

```

DR InterPro; IPR003318; Glyco_hydro_70.
DR Pfam; PF01473; CW_binding_1; 13.
DR Pfam; PF0324; Glyco_hydro_70; 1.
DR TRANSFERASE; Glycosyltransferase; Signal; Repeat; Dental carries.
FT SIGNAL 1 38 POTENTIAL.
FT CHAIN 39 1592 GLUCOSYLTRANSFERASE-1.
FT DOMAIN 39 1044 CATALYTIC (APPROXIMATE).
FT DOMAIN 1093 1592 GLUCAN-BINDING (APPROXIMATE).
FT DOMAIN 1093 1592 6.5 X TANDEM REPEATS.
FT REPEAT 1093 1592 1.
FT REPEAT 1093 1592 2.
FT REPEAT 1158 1207 3.
FT REPEAT 1158 1207 4.
FT REPEAT 1222 1272 5.
FT REPEAT 1287 1337 6.
FT REPEAT 1402 1451 7 (INCOMPLETE).
FT REPEAT 1514 1563 8.
FT REPEAT 1577 1592 9.
SQ SEQUENCE 1592 AA; B0A66D079351ECF CRC64;

Query Match 89.1%; Score 98; DB 1; Length 1592;
Best Local Similarity 86.4%; Pred. No. 2,3e-07;
Matches 19; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 VPSYSPFRAHDSVODLRNII 22
Db 548 VPSYSPFRAHDSVODLRDII 569

RESULT 5
GTF5_STRDO STANDARD; PRT; 1365 AA.
ID GTF5_STRDO
AC P29336;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Glucosyltransferase-S precursor (EC 2.4.1.5) (GTF-S) (Dextranucrase)
DE (Sucrose 6-glucosyltransferase).
OS Streptococcus downei (Streptococcus sobrinus).
DE Streptococcus downei (Streptococcus sobrinus);
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OC NCBI_TaxID=1317;
OK NCBI_TaxID=1317;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MPE28;
RX MEDLINE=90316665; PubMed=2142479;
RA Gilmore K.S., Russell R.R., Ferretti J.J.;
RT "Analysis of the Streptococcus downei gtfS gene, which specifies a
RT glucosyltransferase that synthesizes soluble glucans.";
RL Infect. Immun. 58:2452-2458 (1990).
CC -1- FUNCTION: PRODUCTION OF EXTRACELLULAR GLUCANS, THAT ARE THOUGHT
CC TO PLAY A KEY ROLE IN THE DEVELOPMENT OF THE DENTAL PLAQUE BECAUSE
CC OF THEIR ABILITY TO ADHERE TO SMOOTH SURFACES AND MEDIANE THE
CC AGGREGATION OF BACTERIAL CELLS AND FOOD DEBRIS.
CC -1- CATALYTIC ACTIVITY: Sucrose + ((1,6)-alpha-D-glucosyl)(N) = D-
CC fructose + ((1,6)-alpha-D-glucosyl)(N+1).
CC -1- SUBCELLULAR LOCATION: GLUCAN SYNTHESIS BY GTF-S IS INDEPENDENT OF
CC PRIMER GLUCAN UNLIKE GTF-I.
CC -1- DISEASE: DENTAL CARRIES.
CC -1- MISCELLANEOUS: GTF-S SYNTHESIZES WATER-SOLUBLE GLUCANS (ALPHA
CC 1,6-GLUCOSE).
CC -1- SIMILARITY: TO OTHER GLUCOSYLTRANSFERASES AND SOME TO A GLUCAN-
CC BINDING PROTEIN FROM S.MUTANS.
CC -1- SIMILARITY: Contains 10 cell wall binding repeats.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M30943; AAA26898.1;

```

DR InterPro: IPR002479; CW binding.
DR InterPro: IPR003318; Glyco_hydro_70.
DR Pfam: PF01473; CW_binding_1; 8.
DR Pfam: PF03324; Glyco_hydro_70; 1.
KW Transferase; Glycosyltransferase; Signal; Repeat; Dental caries.
FT SIGNAL 1 36
FT CHAIN 37 1365 GLUCOSYLTRANSFERASE-S.
FT REPEAT 157 177 CELL WALL BINDING 1.
FT REPEAT 178 197 CELL WALL BINDING 2.
FT DOMAIN 198 1061 CATALYTIC (APPROXIMATE).
FT REPEAT 1083 1102 CELL WALL BINDING 3.
FT REPEAT 1083 1102 CELL WALL BINDING 4.
FT REPEAT 1150 1169 CELL WALL BINDING 5.
FT REPEAT 1170 1190 CELL WALL BINDING 6.
FT REPEAT 1225 1243 CELL WALL BINDING 7.
FT REPEAT 1289 1308 CELL WALL BINDING 8.
FT REPEAT 1309 1328 CELL WALL BINDING 9.
FT REPEAT 1331 1352 CELL WALL BINDING 10.
SQ SEQUENCE 1365 AA; 151590 MW; 167296B5A2B8C476 CRC64;

Query Match 69.1%; Score 76; DB 1; Length 1365;
Best Local Similarity 72.7%; Pred. No. 0.0006;
Matches 16; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 VPSYFIRAHDSFVQDLIRNII 22
DB 537 VPMYVIRAHDSFVQRIAKII 558

RESULT 6
GTFD_STRMU STANDARD; PRT; 1462 AA.
AC P49331; 069383; 069386; 069389; 069392; 069398;
DT 01-FEB-1996 (Rel. 33, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Glucosyltransferase-S precursor (EC 2.4.1.5) (GTF-S) (dextranucrase)
DE (Sucrose 6-glucosyltransferase).
GN GTFD OR SMU.910.
OS Streptococcus mutans.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1309;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=GS-5;
RX MEDLINE=91100958; PubMed=2148600;
RA Honda O., Kato C., Kuramitsu H.K.;
RT "Nucleotide sequence of the Streptococcus mutans gtfD gene encoding the glucosyltransferase-S enzyme."
RL J. Gen. Microbiol. 136:2099-2105(1990).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=MT4239 / Serotype C; MT4245 / Serotype E; MT4251 / Serotype F;
RX MT4467 / Serotype B; and MT8148 / Serotype C;
MEDLINE=98231643; PubMed=9570124;
RA Fujiwara T., Terao Y., Hoshino T., Kawabata S., Ooshima T., Sobue S., Kimura S., Hamada S.;
RT "Molecular analyses of glucosyltransferase genes among strains of Streptococcus mutans."
RL FEMS Microbiol. Lett. 161:331-336(1998).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=UAI59 / ATCC 700610 / Serotype C;
RX MEDLINE=22295063; PubMed=12397186;
RA Ajdic D., McShan W.M., McLaughlin R.E., Savic G., Chang J., Carson M.B., Primeaux C., Tian R., Kenton S., Jia H., Lin S., Qian Y., Li S., Zhu H., Najjar F., Lai H., White J., Roe B.A., Ferretti J.J.;
RT "Genome sequence of Streptococcus mutans UAI59, a cariogenic dental pathogen."
RL Proc. Natl. Acad. Sci. U.S.A. 99:14434-14439(2002).
CC -1- FUNCTION: PRODUCTION OF EXTRACELLULAR GLUCANS, THAT ARE THOUGHT TO PLAY A KEY ROLE IN THE DEVELOPMENT OF THE DENTAL PLAQUE BECAUSE

CC OF THEIR ABILITY TO ADHERE TO SMOOTH SURFACES AND MEDIATE THE
CC AGGREGATION OF BACTERIAL CELLS AND FOOD DEBRIS.
CC -1- CATALYTIC ACTIVITY: Sucrose + {(1,6)-alpha-D-glucosyl}(N) = D-fructose + {(1,6)-alpha-D-glucosyl}(N+1).
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- DISEASE: DENTAL CARIES.
CC -1- MISCELLANEOUS: GTF-S SYNTHESIZES WATER-INSOLUBLE GLUCANS (ALPHA 1,3-LINKED GLUCOSE AND SOME 1,6 LINKAGES), GTF-S SYNTHESIZES BOTH WATER-SOLUBLE GLUCANS (ALPHA 1,6-GLUCOSE). GTF-SI SYNTHESIZES BOTH FORMS OF GLUCANS.
CC -1- SIMILARITY: TO OTHER GLUCOSYLTRANSFERASES AND SOME TO A GLUCAN-BINDING PROTEIN FROM S.MUTANS.
CC -1- SIMILARITY: Contains 6 cell wall binding repeats.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M29296; AAA26895.1; -;
CC EMBL; D88653; BAA26103.1; -;
CC EMBL; D88656; BAA26107.1; -;
CC EMBL; D88659; BAA26111.1; -;
CC EMBL; D88652; BAA26115.1; -;
CC EMBL; D89979; BAA26121.1; -;
CC EMBL; AE014932; AAN58619.1; -;
CC InterPro: IPR002479; CW binding.
CC InterPro: IPR003318; Glyco_hydro_70.
CC Pfam: PF03324; Glyco_hydro_70; 1.
KW Transferase; Glycosyltransferase; Signal; Repeat; Dental caries;
KW Complete proteome.
KW FT SIGNAL 1 ?
KW FT CHAIN ? 1462
KW FT DOMAIN 1232 1423
KW FT REPEAT 1232 1295
KW FT REPEAT 1296 1359
KW FT REPEAT 1360 1423
KW FT REPEAT 1360 1423
KW FT VARIANT 10
KW FT VARIANT 19
KW FT VARIANT 58
KW FT VARIANT 68
KW FT VARIANT 81
KW FT VARIANT 113
KW FT VARIANT 122
KW FT VARIANT 132
KW FT VARIANT 135
KW FT VARIANT 137
KW FT VARIANT 202
KW FT VARIANT 255
KW FT VARIANT 275
KW FT VARIANT 288
KW FT VARIANT 301
KW FT VARIANT 313
KW FT VARIANT 317
KW FT VARIANT 328
KW FT VARIANT 350
KW FT VARIANT 628
KW FT VARIANT 688
KW FT VARIANT 726
KW FT VARIANT 730
KW FT VARIANT 762
KW FT VARIANT 762
KW POTENTIAL.
KW GLUCOSYLTRANSFERASE-S
KW 3 X 63 AA APPROXIMATE TANDEM REPEATS.
KW 1.
KW 2.
KW 3.
KW Y -> H (IN STRAINS GS-5, MT4239, MT4245, MT4251, MT4467 AND MT8148).
KW I -> V (IN STRAINS GS-5, MT4239, MT4245, MT4251, MT4467 AND MT8148).
KW K -> E (IN STRAIN MT4467).
KW A -> S (IN STRAINS MT4239 AND MT4245).
KW A -> T (IN STRAINS MT4251 AND MT8148).
KW T -> I (IN STRAINS MT4239 AND MT4245).
KW A -> V (IN STRAINS MT4239, MT4245 AND MT8148).
KW A -> S (IN STRAINS GS-5 AND MT4467).
KW D -> N (IN STRAIN MT4245).
KW A -> T (IN STRAINS GS-5, MT4239, MT4245, MT4251, MT4467 AND MT8148).
KW V -> L (IN STRAIN MT4239).
KW D -> N (IN STRAIN MT8148).
KW E -> D (IN STRAINS MT4239, MT4245 AND MT4251).
KW D -> N (IN STRAINS MT4239, MT4245 AND MT4251).
KW Q -> H (IN STRAIN MT4245).
KW D -> N (IN STRAINS MT4239 AND MT4251).
KW E -> K (IN STRAIN MT4239).
KW V -> F (IN STRAIN MT4239).
KW F -> L (IN STRAINS MT4239, MT4251 AND MT4467).
KW KKKTQ -> EKEVTL (IN STRAIN MT4251).
KW A -> S (IN STRAIN MT4239).
KW IDOGSEA -> ADKGNDS (IN STRAIN MT4251).
KW IDOGS -> ADKGN (IN STRAINS MT4239 AND MT4245).
KW T -> A (IN STRAINS GS-5, MT4239, MT4245, MT4251, MT4467 AND MT8148).


```

FT  VARIANT 964 964 D -> Y (IN STRAIN MT4251).
FT  VARIANT 1019 1019 B -> K (IN STRAINS MT4245 AND MT4251).
FT  VARIANT 1059 1059 LG -> IR (IN STRAIN MT4251).
FT  VARIANT 1060 1060 G -> R (IN STRAIN MT4245).
FT  VARIANT 1080 1080 G -> R (IN STRAIN MT4239).
FT  VARIANT 1142 1142 H -> Q (IN STRAIN GS-5).
FT  VARIANT 1198 1198 S -> N (IN STRAIN MT4239).
FT  VARIANT 1220 1220 Y -> C (IN STRAINS MT4251 AND MT4467).
FT  VARIANT 1280 1280 F -> L (IN STRAIN MT4467).
FT  VARIANT 1282 1282 Q -> P (IN STRAIN MT4245).
FT  VARIANT 1290 1290 K -> T (IN STRAIN MT4245).
FT  VARIANT 1311 1311 N -> D (IN STRAIN MT4245).
FT  VARIANT 1403 1403 G -> D (IN STRAINS GS-5 AND MT4467).
FT  VARIANT 1425 1425 G -> R (IN STRAIN GS-5).
FT  VARIANT 1449 1449 R -> K (IN STRAIN MT4467).
FT  CONFLICT 1428 1462 R (IN REF. 1).
SQ  SEQUENCE 1462 AA; 16387 MW; CE4A279C4D708645 CRC64;
Query Match 61.8%; Score 68; DB 1; Length 1462;
Best Local Similarity 70.0%; Pred. No. 0.012; Mismatches 4; Indels 0; Gaps 0;
Matches 14; Conservative 2;
Oy 3 SYSFIRAHSEVQDLIRNI 22
Db 576 NYFIRAHSEVQVYIAKI 595

```

```

CC  Name=2;
CC  IsoId=P17342-2; Sequence=VSP 001812;
CC  -1- MISCELLANEOUS: THESE SEEM TO BE AT LEAST THREE ANP RECEPTORS: TWO
CC  WITH GUANYLATE CYCLASE ACTIVITY (ANP-A AND ANP-B) AND ONE (ANP-C)
CC  WHICH IS PROBABLY RESPONSIBLE FOR THE CLEARANCE OF ANP FROM THE
CC  CIRCULATION WITHOUT A ROLE IN SIGNAL TRANSDUCTION.
CC  -1- SIMILARITY: TO ANP-A AND ANP-B RECEPTORS IN THEIR EXTRACELLULAR
CC  AND TRANSMEMBRANE DOMAINS.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
CC  EMBL: X52282; CAA36523.1; -
CC  EMBL: M59305; AAA51734.1; -
CC  EMBL: AF025998; AAB88801.1; -
CC  PIR: S10150; OYHUCR.
CC  PIR: JUDN; 05-SEP-01.
CC  Genew; HGNC:7945; NPR3.
CC  MIM: 108962; -
CC  GO: GO:0004872; F:receptor activity; TAS.
CC  GO: GO:0001501; P:skeletal development; TAS.
CC  InterPro; IPR001828; ANP_receptor.
CC  InterPro; IPR001170; Ntpep_receptor.
CC  Pfam; PF01094; ANP_receptor; 1.
CC  PRINTS; PS00458; ANP_RECEPTORS; 1.
CC  Prosite; PS00458; ANP_RECEPTORS; 1.
CC  Receptor; Glycoprotein, Transmembrane, Signal, Alternative splicing;
CC  3D-structure.
CC  3D-SIGNAL 1 26
CC  FT CHAIN 27 541
CC  FT DOMAIN 27 481
CC  FT TRANSMEM 482 504
CC  FT DOMAIN 505 541
CC  FT DISULFID 108 136
CC  FT DISULFID 213 261
CC  FT DISULFID 473 473
CC  FT CARBOHYD 86 86
CC  FT CARBOHYD 293 293
CC  FT CARBOHYD 394 394
CC  FT VARSPLIC 476 477
CC  SQ SEQUENCE 541 AA; 59807 MW; 8A66415F7F7D62B7 CRC64;
Query Match 43.6%; Score 48; DB 1; Length 541;
Best Local Similarity 44.4%; Pred. No. 6.2; Mismatches 5; Indels 0; Gaps 0;
Matches 8; Conservative 5;
Oy 4 YSFIRAHSEVQDLIRNI 21
Db 234 YSFDETKDLDEIVANI 251

```

```

RESULT 8
MAD2 YEAST
ID MAD2_YEAST STANDARD; PRT; 196 AA.
AC P40958;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Mitotic spindle checkpoint component MAD2 (Mitotic MAD2 protein).
GN MAD2 OR YJL030W OR YJ156.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.

```

[illegible]

[1]
RN SEQUENCE FROM N.A.
RP MEDLINE=92199240; PubMed=1802034;
RX Genhaufte F.S., Cooper T.G.;
RA "The urea amidolyase (DURL2) gene of Saccharomyces cerevisiae";
RT DNA Seq. 2:19-32(1991).
RL
RN [2]
RP SEQUENCE OF 1-893 FROM N.A.
RC STRAIN=S288c;
RA Rieger M.;
RL Submitted (AUG-1994) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE OF 873-1835 FROM N.A.
RC STRAIN=S288c;
RA Feldmann H., Mannheim G., Schwarzlouse C., Vetter I.;
RL Submitted (AUG-1994) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE OF 1487-1835 FROM N.A.
RC STRAIN=S288c;
RX Buseigneur F., Mallet L., Gallon L., Uacquet M.;
RA "A 12.8 kb segment, on the right arm of chromosome II from
RT Saccharomyces cerevisiae including part of the DURL2 gene, contains
RT five putative new genes."
Yeast 9:797-806(1993)
CC - FUNCTION: HYDROLYSIS OF UREA TO AMMONIA AND CO2.
CC - CATALYTIC ACTIVITY: ATP + urea + CO(2) = ADP + phosphate + urea-l-
CC carboxylate.
CC - CARBOLYTIC ACTIVITY: Urea-l-carboxylate + H(2)O = 2 CO(2) + 2
CC NH(3).
CC - COPACTOR: BIOTIN.
CC - PATHWAY: ALLANTOIN AND ARGININE METABOLISM.
CC - SUBUNIT: Monomer.
CC - INDUCTION: BY ALLOPHANATE OR ITS NON-METABOLIZED ANALOG
CC OXALURATE. RERESSED IN THE PRESENCE OF READILY USED NITROGEN
CC SOURCES.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/) announce/
CC or send an email to license@isb-sib.ch).
CC -----

DR EMBL; MG4926; AAC1463.1; -
DR EMBL; Z36077; CAA85172.1; -
DR EMBL; Z21487; CAA79695.1; -
DR PIR; S46082; S46082.
DR HSSP; P24182; IBNC.
DR SGD; S00000412; DURL2.
DR GO; GO:0004039; F:allophanate hydrolase activity; IMP.
DR GO; GO:0004847; F:urea carboxylase activity; IMP.
DR GO; GO:0019627; F:urea metabolism; IMP.
DR InterPro; IPRO00120; Amidase.
DR InterPro; IPRO01882; Biotin attach.
DR InterPro; IPRO005482; Biotin carb C.
DR InterPro; IPRO000089; Biotin lipoyl.
DR InterPro; IPRO05479; CPase_L_D2.
DR InterPro; IPRO05481; CPase_L_N.
DR InterPro; IPRO03778; DUFL83.
DR InterPro; IPRO03833; DUF213.
DR Pfam; PF02682; AHS1; 1.
DR Pfam; PF02626; AHS2; 1.
DR Pfam; PF01422; Amidase; 1.
DR Pfam; PF02785; Biotin carb C; 1.
DR Pfam; PF00364; Biotin lipoyl; 1.
DR Pfam; PF00289; CPSase_L_chain; 1.
DR Pfam; PF02786; CPSase_L_D2; 1.
DR PROSITE; PS00188; BIOTIN; 1.
DR PROSITE; PS00865; CPSASE_1; 1.
DR PROSITE; PS00867; CPSASE_2; 1.
KW ligase; Hydrolase; Multifunctional enzyme; Arginine metabolism;

KW Biotin: ATP-binding. ATP (POTENTIAL).
 FT NP_BIND 122 129
 FT BINDING 1798 1798
 FT CONFLICT 96 96
 FT CONFLICT 256 258
 FT CONFLICT 459 459
 FT CONFLICT 830 830
 FT CONFLICT 1395 1395
 FT SEQUENCE 1835 AA; P52BDDDFE42CD65 CRC64;
 SQ
 Query Match 40.9%; Score 45; DB 1; Length 1835;
 Best Local Similarity 42.1%; Pred. No. 70;
 Matches 8; Conservative 7; Mismatches 4; Indels 0; Gaps 0;

QY 2 PSYSFIRADSEVODLRN 20
 DB 102 PSFAYEFSKSKYVELLRN 120

RESULT 10
 BACC_BACLI STANDARD; PRT; 6359 AA.
 ID BACC_BACLI
 AC 068008;
 DT 15-JUL-1999 (Rel. 38, Created)
 DT 15-JUL-1999 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Bactracin synthetase 3 (BA3) [includes: ATP-dependent isoleucine
 adenylation (IleA) (isoleucine acetylase); ATP-dependent histidine
 adenylation (HisA) (histidine acetylase); ATP-dependent aspartate
 adenylation (D-AspA) (D-aspartate acetylase); Aspartate racemase
 adenylation (AspA) (aspartate acetylase); Aspartate racemase
 (EC 5.1.1.13); Phenylalanine racemase [ATP hydrolyzing]
 (EC 5.1.1.11)].
 DE BACC.
 GN Bacillus licheniformis.
 OS Bacillus licheniformis.
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 OK NCBI_TaxId=1402;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 10716;
 RX MEDLINE=98089193; PubMed=9427658;
 RA Konz D., Krens A., Schoergerdoffer K., Marahiel M.A.;
 RT "The bacitracin biosynthesis operon of Bacillus licheniformis ATCC
 10716: molecular characterization of three multi-modular peptide
 synthetases.";
 RT Chem. Biol. 4:927-937(1997).
 RL
 CC -1- FUNCTION: INDUCES PEPTIDE SYNTHESIS, ACTIVATES AND INCORPORATES
 FIVE AMINO ACIDS, FORMS A THIAZOLINE RING BETWEEN THE FIRST TWO
 AMINO ACIDS AND INCORPORATES A D-GLUTAMINE IN THE FOURTH POSITION.
 CC -1- CATALYTIC ACTIVITY: L-aspartate = D-aspartate.
 CC -1- CATALYTIC ACTIVITY: ATP + L-phenylalanine = AMP + diphosphate + D-
 phenylalanine.
 CC -1- COFACTOR: CONTAINS 5 COVALENTLY BOUND PHOSPHOPANTETHEINES
 (POTENTIAL).
 CC -1- PATHWAY: Cyclic peptide antibiotic bacitracin biosynthesis.
 CC -1- SUBUNIT: LARGE MULTISUBUNIT COMPLEX OF BAI, BAI2 AND BAI3.
 CC -1- DOMAIN: CONSISTS OF FIVE MODULES WITH TWO EPIMERIZATION DOMAINS IN
 THE SECOND AND FOURTH MODULES, AND A PUTATIVE C-TERMINAL
 THIOESTERASE DOMAIN. EACH MODULE INCORPORATES ONE AMINO ACID INTO
 THE PEPTIDE PRODUCT AND CAN BE FURTHER SUBDIVIDED INTO DOMAINS
 RESPONSIBLE FOR SUBSTRATE ADENYLATION, THIOALATION, CONDENSATION
 (NOT FOR THE INITIATION MODULE), AND EPIMERIZATION (OPTIONAL), AND
 N-METHYLATION (OPTIONAL).
 CC -1- MISCELLANEOUS: BACTIRACIN IS A MIXTURE OF AT LEAST TEN CYCLIC
 DODECAPETIDES, THAT DIFFER BY ONE OR TWO AMINO ACIDS. THE MOST
 ABUNDANT IS BACTIRACIN A, A BRANCHED CYCLIC DODECAPETIDE IT
 CONTAINS AN N-TERMINAL LINEAR PENTAPEPTIDE MOIETY (Ile-Cys-Ileu-D-
 Glu-Ile) WITH AN ISOLEUCINE-CYSTEINE THIAZOLINE CONDENSATION
 PRODUCT AND A C-TERMINAL HEPTAPEPTIDE RING (Iys-D-Orn-Ile-D-Phe-
 His-D-Asp-Asn), IN WHICH THE PHE ALPHA-CARBOXY GROUP OF THE C-
 TERMINAL ASN IS BOUND TO THE FIRST-AMINO GROUP OF LEUCINE. IT
 CONTAINS FOUR AMINO ACIDS IN THE D-CONFIGURATION (GLU-4, ORN-7,

CC PHE-9, AND ASP-11).
 CC -1- SIMILARITY: BELONGS TO THE ATP-DEPENDENT AMP-BINDING ENZYME
 FAMILY.
 CC -1- SIMILARITY: CONTAINS 5 acyl carrier domains.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: AF007865; AAC06348.1; -
 DR PIR: T31679; T31679.
 DR HSSP: P14687; IANU.
 DR InterPro: IPR000873; AMP-bind.
 DR InterPro: IPR001242; Condensatn.
 DR InterPro: IPR006163; Pp_bind.
 DR InterPro: IPR006162; Ppantne_attach.
 DR InterPro: IPR001031; Thioesterase.
 DR InterPro: AMP-binding; 5.
 DR Pfam: PF00668; Condensation; 7.
 DR Pfam: PF00550; Pp-binding; 5.
 DR Pfam: PF00975; Thioesterase; 1.
 DR PROSITE: PS00012; PHOSPHOPANTETHEINE; 4.
 DR PROSITE: PS00455; AMP BINDING; 5.
 DR PROSITE: PS00075; ACP DOMAIN; 5.
 KW Lysase; isomerase; Hydrolase; Multifunctional enzyme; Repeat.
 KW Phosphopantetheine; Multifunctional enzyme; Repeat.
 FT REPEAT 461 1034
 FT REPEAT 1517 2064
 FT REPEAT 2999 3570
 FT REPEAT 4047 4612
 FT REPEAT 5549 6129
 FT DOMAIN 966 1033
 FT DOMAIN 1998 2064
 FT DOMAIN 3502 3569
 FT DOMAIN 4544 4610
 FT DOMAIN 6052 6119
 FT BINDING 996 996
 FT BINDING 2028 2028
 FT BINDING 3532 3532
 FT BINDING 4574 4574
 FT BINDING 6082 6082
 SQ SEQUENCE 6359 AA; 722923 MW; 82A273C546253074 CRC64;
 Query Match 40.9%; Score 45; DB 1; Length 6359;
 Best Local Similarity 45.5%; Pred. No. 2.7e+02;
 Matches 10; Conservative 4; Mismatches 4; Indels 4; Gaps 1;

QY 1 VPSYFIRADSEVODLRNII 22
 DB 3186 VPSFSF---DSVDFITTLI 3203

RESULT 11
 YIAU_ECOLI STANDARD; PRT; 282 AA.
 ID YIAU_ECOLI
 AC P37671;
 DT 01-OCT-1994 (Rel. 30, Created)
 DT 01-OCT-1994 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 30, Last annotation update)
 DE Hypothetical transcriptional regulator yiaU.
 DE YIAU OR B3574.
 GN Escherichia coli.
 OS Escherichia coli.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Escherichia.
 OK NCBI_TaxId=562;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=K12 / MG1655;
 RX MEDLINE=94316500; PubMed=8041620;

RA Sofia H.J., Burland V., Daniels D.L., Plunkett G. III, Blattner F.R.;
 RT "Analysis of the Escherichia coli genome. V. DNA sequence of the
 RT region from 76.0 to 81.5 minutes.";
 RL Nucleic Acids Res. 22:2576-2586(1994).
 CC -1- SIMILARITY: BELONGS TO THE ICLR FAMILY OF TRANSCRIPTIONAL
 CC REGULATORS. STRONG, TO H.INFLUENZAE H11032.
 CC -----
 CC THE SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: U00039; AAB18551.1; -
 CC EMBL: AE000435; AAC76598.1; -
 CC PIR: S47795; S47795.
 CC Ecogene: EG12278; YIAU.
 CC InterPro: IPR005471; HTH_ICLR.
 CC InterPro: IPR005473; HTH_ICLRlike.
 CC Pfam: PF01614; ICLR; 1.
 CC SMART: SM00346; HTH_ICLR; 1.
 CC PROSITE: PS01051; HTH_ICLR_FAMILY; 1.
 CC DR Hypothetical protein; Transcription regulation; DNA-binding;
 CC KW Complete proteome.
 CC FT DNA BIND 45 64 H-T-H MOTIF (POTENTIAL).
 CC SEQUENCE 282 AA; 31066 MW; 228887B672B958F CRC64;
 SQ
 Query Match 40.0%; Score 44; DB 1; Length 282;
 Best Local Similarity 43.5%; Pred. No. 13;
 Matches 10; Conservative 3; Mismatches 8; Indels 2; Gaps 1;
 QY 2 PSY-SFIRAHSEVODLIRNTI 22
 DB 169 PDVKSVMESHQHEIPLRNTI 191
 RESULT 12
 M12D_BACSU STANDARD; PRT; 344 AA.
 AC P26935;
 DT 01-AUG-1992 (Rel. 23, Created)
 DT 01-AUG-1992 (Rel. 23, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Myo-inositol 2-dehydrogenase (EC 1.1.1.18).
 GN IDH OR IDLG OR E83G.
 OS Bacillus subtilis.
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 OX NCB1_TaxID=1423;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=168 / 60015;
 RA MEDLINE=92104493; PubMed=1761221.
 RA Fujita Y., Shindo K., Miwa Y., Yoshida K.,
 RT "Bacillus subtilis inositol dehydrogenase-encoding gene (idh):
 RT sequence and expression in Escherichia coli.";
 RL Gene 108:121-125(1991).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=168 / BGSC1A1;
 RA MEDLINE=95039891; PubMed=7952181;
 RA Yoshida K., Sano H., Miwa Y., Ogasawara N., Fujita Y.,
 RT "Cloning and nucleotide sequencing of a 15 kb region of the Bacillus
 RT subtilis genome containing the idh operon.";
 RL Microbiology 140:2289-2298(1994).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=168;
 RA MEDLINE=98044033; PubMed=9384377;
 RA Kunze F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
 RA Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,
 RA Boriss R., Bourstier L., Brans A., Braun M., Brignell S.C., Bron S.,

RA Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,
 RA Choi S.K., Codani J.M., Conneron I.F., Cummings N.J., Daniel R.A.,
 RA Danio F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emerson P.T.,
 RA Entian K.D., Errington J., Fabre C., Ferrari E., Foulger D.,
 RA Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Gallier N.,
 RA Gilm S.Y., Glaeser P., Gottreau A., Goughly E.J., Grandi G.,
 RA Giuseppe G., Guy B.J., Haga K., Halech U., Harwood C.R., Henaut A.,
 RA Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,
 RA Joris B., Karamata D., Kasahara Y., Kleer-Blanchard M., Klein C.,
 RA Kobayashi Y., Koetter P., Koningsstein G., Krogh S., Kumano M.,
 RA Kunita K., Lapius A., Lardinois S., Lauber J., Lazarevic V.,
 RA Lee S.M., Levine A., Liu H., Masuda S., Manel C., Medigue C.,
 RA Medina N., Meliardo R.P., Mizuno M., Moestl D., Nakai S., Noback M.,
 RA Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,
 RA Parro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,
 RA Prescan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,
 RA Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadate Y.,
 RA Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,
 RA Sekiguchi J., Sekowska A., Seror S.J., Serron F., Shin B.S., Soldo B.,
 RA Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru A.,
 RA Takeuchi M., Tamakoshi A., Tanaka T., Terpestra P., Tognoni A.,
 RA Tozato V., Uchiyama S., Vandenbol M., Vannier F., Vasariotti A.,
 RA Viart A., Wandut R., Wedler E., Wedler H., Wellzner T.,
 RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
 RA Yoshida K., Yoshikawa H.F., Zumbstein E., Yoshikawa H., Danchin A.,
 RT "The complete genome sequence of the Gram-positive bacterium Bacillus
 RT subtilis".
 RT Nature 390:249-256(1997).
 RN [4]
 RP CHARACTERIZATION.
 RA MEDLINE=79239346; PubMed=112095;
 RA Ramaley R., Fujita Y., Freese E.;
 RT "Purification and properties of Bacillus subtilis inositol
 RT dehydrogenase.";
 RL J. Biol. Chem. 254:7684-7690(1979).
 CC -1- CATALYTIC ACTIVITY: Myo-inositol + NAD(+) = 2,4,6/3,5-
 CC pentahydroxycyclohexanone + NADH.
 CC -1- PATHWAY: Myo-inositol catabolism; first step.
 CC -1- SUBUNIT: Monomer.
 CC -1- INDUCTION: INOSITOL, SUBJECTED TO CATABOLITE REPRESSION.
 CC -1- SIMILARITY: BELONGS TO THE GFO/IDH/MOCA FAMILY.
 CC -----
 CC THE SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: M76431; AAA22543.1; -
 CC EMBL: D14399; BAA03296.1; -
 CC EMBL: Z99124; CAB16006.1; -
 CC PIR: JH0511; JH0511; idh.
 CC Subtilist; BG10693; idh.
 CC InterPro: IPR000683; GFO_IDH_MOCA.
 CC InterPro: IPR004104; GFO_IDH_MOCA_C.
 CC Pfam: PF01408; GFO_IDH_MOCA; 1.
 CC Pfam: PF02894; GFO_IDH_MOCA_C; 1.
 CC DR Oxidoreductase, NAD, Complete proteome.
 CC SEQUENCE 344 AA; 38351 MW; 2FCB508D4E2C332F CRC64;
 SQ
 Query Match 40.0%; Score 44; DB 1; Length 344;
 Best Local Similarity 50.0%; Pred. No. 16;
 Matches 8; Conservative 4; Mismatches 4; Indels 0; Gaps 0;
 QY 6 FIRAHSEVODLIRNTI 21
 DB 276 FVAAYDVEIOPFIDSI 291
 RESULT 13
 YNC4_YEAST

```

ID YN14 YEAST STANDARD; PRT; 587 AA.
AC P53933.
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Hypothetical 66.1 kDa protein in RPS7B-YPT53 intergenic region.
GN YNL094W OR N2219.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycaceae;
OC NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=8288c / FY1679;
RC MEDLINE=96367601; PubMed=8771715;
RA Garcia-Camalejo J.M., Boskovic U., Jimenez A.;
RT "Sequence analysis of a 14.2 kb fragment of Saccharomyces cerevisiae
RT chromosome XIV that includes the YPT53, CRN150 and gsr m2 genes and
RT four new open reading frames.";
RL Yeast 12:599-608(1996).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X85811; CAAS9823.1; ALT_INT.
DR EMBL; Z71370; CA95970.1; -.
DR PIR; S63033; S63033.
DR SGD; S0005038; YNL094W.
DR GO; GO:0015629; C:actin cytoskeleton organization and biogenesis; IPI.
DR GO; GO:0030036; P:actin cytoskeleton organization and biogenesis; IPI.
DR Hypothetical protein.
KW SEQUENCE 587 AA; 66134 MW; B85C525548BA34BC CRC64;
SQ
Query Match 40.0%; Score 44; DB 1; Length 587;
Best Local Similarity 45.0%; Pred. No. 29;
Matches 9; Conservative 4; Mismatches 7; Indels 0; Gaps 0;
OY 3 SYSFIRANDSEYVDLIRNII 22
Db 430 SMSDPVSHDEYVNEVNNII 449

```

```

CC -1- FUNCTION: PHOSPHORYLATES PTDINS. PTDINS4P AND PTDINS(4,5)P2 WITH A
CC PREFERENCE FOR PTDINS(4,5)P2.
CC -1- CATALYTIC ACTIVITY: ATP + 1-phosphatidyl-1D-myo-inositol 4,5-
CC bisphosphate = ADP + 1-phosphatidyl-1D-myo-inositol 3,4,5-
CC trisphosphate.
CC -1- SUBUNIT: HETERODIMER OF A P110 (CATALYTIC) AND A P85 (REGULATORY)
CC SUBUNIT.
CC -1- SIMILARITY: BELONGS TO THE PI3/P14-KINASES FAMILY.
CC -1- SIMILARITY: Contains 1 C2 domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M93252; AAA30698.1; -.
DR PIR; A43322; A43322.
DR InterPro; IPR000008; C2.
DR InterPro; IPR000403; PI3_P14_Kinase.
DR InterPro; IPR002420; PI3K_C2.
DR InterPro; IPR003113; PI3K_P85B.
DR InterPro; IPR000341; PI3K_ras_Bind.
DR InterPro; IPR01263; PI3Ka.
DR Pfam; PF00454; PI3_P14_Kinase; 1.
DR Pfam; PF00792; PI3K_C2; 1.
DR Pfam; PF02192; PI3K_P85B; 1.
DR Pfam; PF00794; PI3K_P85; 1.
DR Pfam; PF00613; PI3Ka; 1.
DR SMART; SM00239; C2; 1.
DR SMART; SM00142; PI3K_C2; 1.
DR SMART; SM00143; PI3K_P85B; 1.
DR SMART; SM00144; PI3K_P85; 1.
DR SMART; SM00145; PI3Ka; 1.
DR SMART; SM00146; PI3K; 1.
DR PROSITE; PSS0004; C2 DOMAIN 2; FALSE_NEG.
DR PROSITE; PSS00915; PI3_4_KINASE_1; 1.
DR PROSITE; PSS00916; PI3_4_KINASE_2; 1.
DR PROSITE; PSS0290; PI3_4_KINASE_3; 1.
DR PROSITE; PSS0290; PI3_4_KINASE_3; 1.
DR TRANSFERASE; Kinase; Multigene family.
KW DOMAIN 319 428 PI3K/PI4K
FT DOMAIN 797 1068 PI3K/PI4K
SQ SEQUENCE 1068 AA; 124327 MW; C753DC2F39FDDF0 CRC64;
Query Match 40.0%; Score 44; DB 1; Length 1068;
Best Local Similarity 45.5%; Pred. No. 56;
Matches 10; Conservative 2; Mismatches 10; Indels 0; Gaps 0;
OY 1 VPSYFIRANDSEYVDLIRNII 22
Db 123 MPVCEFDVXDPVEVDFFRNII 144

```

```

RESULT 14
P11A_BOVIN STANDARD; PRT; 1068 AA.
ID P11A_BOVIN
AC P32871.
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Phosphatidylinositol-4,5-bisphosphate 3-kinase catalytic subunit.
DE alpha isoform (EC 2.7.1.153) (PI3-kinase p110 subunit alpha) (PtdIns-
DE 3-kinase p110) (PI3K).
GN PIK3CA.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OC NCBI_TaxID=9913;
OX [1]
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC TISSUE=Brain;
RC MEDLINE=92354059; PubMed=1322797;
RA Hiles I.D., Otsu M., Volinia S., Fry M.J., Gout I., Dhand R.,
RA Panayotou G., Ruiz-Larrea F., Thompson A., Totty N.F., Hsuan J.J.,
RA Courtneidge S.A., Parker P.J., Waterfield M.D.;
RT "Phosphatidylinositol 3-kinase: structure and expression of the 110
RT kd catalytic subunit.";
RL Cell 70:419-429(1992).

```

```

RESULT 15
P11A_HUMAN STANDARD; PRT; 1068 AA.
ID P11A_HUMAN
AC P42336.
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Phosphatidylinositol-4,5-bisphosphate 3-kinase catalytic subunit.
DE alpha isoform (EC 2.7.1.153) (PI3-kinase p110 subunit alpha) (PtdIns-
DE 3-kinase p110) (PI3K).
GN PIK3CA.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OC NCBI_TaxID=9606;
OX [1]
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=95229146; PubMed=7713498;

```


GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd

OM protein - protein search, using sw model

Run on: November 13, 2003, 09:31:40 ; Search time 33.1564 Seconds
(without alignments)
171.224 Million cell updates/sec

```
Title: US-09-290-049A-15
Perfect score: 110
Sequence: 1 VPSYSFIRAHDSVQDLIRNII 22
```

Scoring table:	BLOSUM62	
	Gapopen 10	Gapext 0.5

```
Searched:      830525 seqs, 258052604 residues
Total number of hits satisfying chosen parameters:  830525
```

```
Minimum DB seq length: 0
Maximum DB seq length: 20000000000
```

Post-processing:	Minimum Match	0%
	Maximum Match	100%
	Listing first	45 summaries

```

1:  sp archaea:*
2:  sp bacteria:*
3:  sp fungi:*
4:  sp human:*
5:  sp invertebrate:*
6:  sp mammal:*
7:  sp mhc:*
8:  sp organelle:*
9:  sp phage:*
10: sp plant:*
11: sp rodent:*
12: sp virus:*
13: sp vertebrate:*
14: sp unclassified:*
15: sp virus:*
16: sp bacteriap:*
17: sp archaeap:*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

Result	No.	Score	Query Match	Length	DB	ID	Description
1	98	89.1	1590	2	059983	059983	streptococc
2	98	89.1	1590	2	055263	055263	streptococc
3	83	75.5	1290	2	048756	048756	leuonostoc
4	83	75.5	1477	2	091466	091466	leuonostoc
5	83	75.5	1508	2	09E2H5	09e2h5	leuonostoc
6	83	75.5	1508	2	052224	052224	leuonostoc
7	82	74.5	1527	2	092AR4	092ar4	leuonostoc
8	82	74.5	1527	2	08KXH1	08kxh1	leuonostoc
9	82	74.5	1554	2	08KZL5	08kzl5	streptococc
10	80	72.7	1016	2	091CJ7	091cj7	leuonostoc
11	78	70.9	2835	2	08G9Q2	08g9q2	leuonostoc
12	77	70.0	1512	2	09MXJ5	09wxj5	streptococc
13	76	69.1	1338	2	09MXJ4	09wxj4	streptococc
14	71	64.5	1449	2	068542	068542	streptococc
15	71	64.5	1449	2	055264	055264	streptococc
16	71	64.5	2057	2	09RE05	09re05	leuonostoc

17	70	63.6	1577	2	Q552S5	Q552S5 streptococc
18	68	61.8	1518	2	Q06000	Q06000 streptococc
19	68	61.8	1575	2	Q9LCH3	Q9LCH3 streptococc
20	68	61.8	1577	2	Q54178	Q54178 streptococc
21	68	61.8	1599	2	Q00599	Q00599 streptococc
22	50	45.5	51	16	Q9RNP7	Q9RNP7 vibrio cho
23	48	43.6	93	2	Q92IX3	Q92IX3 borrelia bu
24	48	43.6	244	17	Q8Q0D3	Q8Q0D3 methanosar
25	47	42.7	480	5	Q9SMU6	Q9SMU6 giardia lam
26	47	42.7	639	2	Q9AKQ3	Q9AKQ3 bacillus sp
27	47	42.7	1534	3	Q8U0Z6	Q8U0Z6 cryptococcu
28	47	42.7	1567	3	Q8U0W2	Q8U0W2 cryptococcu
29	46.5	42.3	114	10	Q65491	Q65491 arabidopsi
30	46	41.8	66	2	Q9KXK6	Q9KXK6 bacillus me
31	46	41.8	200	16	Q9KXJ7	Q9KXJ7 clostridiu
32	46	41.8	326	12	Q8QVU6	Q8QVU6 avian reov
33	46	41.8	597	16	Q9PR58	Q9PR58 ureaplasma
34	46	41.8	631	16	Q98PL8	Q98PL8 mycoplasma
35	46	41.8	1032	10	Q9FJY9	Q9FJY9 arabidops
36	46	41.8	1035	10	Q8GS71	Q8GS71 arabidops
37	45.5	41.4	248	16	Q987H1	Q987H1 rhizobium
38	45.5	41.4	767	3	Q74759	Q74759 schizosacc
39	45.5	41.4	956	16	Q9PFU0	Q9PFU0 chlamydia
40	45	40.9	111	16	Q8XLB3	Q8XLB3 clostridi
41	45	40.9	136	9	Q64072	Q64072 bacterioph
42	45	40.9	136	16	Q31949	Q31949 bacillus s
43	45	40.9	326	12	Q8QV08	Q8QV08 avian reov
44	45	40.9	852	13	Q9PTES	Q9PTES xenopus l
45	44.5	40.5	462	2	Q8GTY5	Q8GTY5 mycoplasma

ALIGNMENTS

RESULT 1				
059983				
ID	059983	PRELIMINARY;	PRT;	1590 AA.
AC	059983			
AD	059983			
DT	01-NOV-1996 (TrEMBLrel. 01, Created)			
DT	01-NOV-1996 (TrEMBLrel. 01, Last sequence update)			
DT	01-OCT-2002 (TrEMBLrel. 22, Last annotation update)			
DE	Glucosyltransferase-I precursor (EC 2.4.1.5).			
GN	GTFI.			
OS	Streptococcus sobrinus.			
OC	Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;			
OC	Streptococcus.			
OX	NCBI_Taxid=1310;			
RN	[1]			
RP	SEQUENCE FROM N.A..			
RC	STRAIN=OM176;			
RC	MEDLINE=94146405; PubMed=8312602;			
RA	Sato S., Inoue M., Handa N., Aizawa Y., Katayama T.;			
RT	"DNA sequence of the glucosyltransferase gene of serotype d			
RT	Streptococcus sobrinus.";			
RL	DNA Seq. 4:19-27(1993).			
DR	EMBL; D13858; BAA02976.1; -;			
DR	InterPro; IPR002479; CM_binding.			
DR	InterPro; IPR00318; Glyco_hydro.70.			
DR	Pfam; PF01473; CM_binding.1; 16.			
DR	Pfam; PF02324; Glyco_hydro.70; 1.			
KM	Glycosyltransferase; Signal; transferase.			
FT	SIGNAL			
FT	CHAIN			
FT	SEQUENCE			
FT	1590 AA;			
FT	175955 NM;			
FT	C3C83A57CF3C280E CRC64;			
FT	39			
FT	1590			
FT	GLUCOSYLTRANSFERASE-I.			
FT	POTENTIAL.			
FT	1			
FT	39			
FT	1590			
FT	175955 NM;			
FT	C3C83A57CF3C280E CRC64;			
FT	39			
FT	1590			
FT	GLUCOSYLTRANSFERASE-I.			
FT	POTENTIAL.			
FT	1			
FT	39			
FT	1590			
FT	175955 NM;			
FT	C3C83A57CF3C280E CRC64;			
FT	39			
FT	1590			
FT	GLUCOSYLTRANSFERASE-I.			
FT	POTENTIAL.			
FT	1			
FT	39			
FT	1590			
FT	175955 NM;			
FT	C3C83A57CF3C280E CRC64;			
FT	39			
FT	1590			
FT	GLUCOSYLTRANSFERASE-I.			
FT	POTENTIAL.			
FT	1			
FT	39			
FT	1590			
FT	175955 NM;			
FT	C3C83A57CF3C280E CRC64;			
FT	39			
FT	1590			
FT	GLUCOSYLTRANSFERASE-I.			
FT	POTENTIAL.			
FT	1			
FT	39			
FT	1590			
FT	175955 NM;			
FT	C3C83A57CF3C280E CRC64;			
FT	39			
FT	1590			
FT	GLUCOSYLTRANSFERASE-I.			
FT	POTENTIAL.			
FT	1			
FT	39			
FT	1590			
FT	175955 NM;			
FT	C3C83A57CF3C280E CRC64;			
FT	39			
FT	1590			
FT	GLUCOSYLTRANSFERASE-I.			
FT	POTENTIAL.			
FT	1			
FT	39			
FT	1590			
FT	175955 NM;			
FT	C3C83A57CF3C280E CRC64;			
FT	39			
FT	1590			
FT	GLUCOSYLTRANSFERASE-I.			
FT	POTENTIAL.			
FT	1			
FT	39			
FT	1590			
FT	175955 NM;			
FT	C3C83A57CF3C280E CRC64;			
FT	39			
FT	1590			
FT	GLUCOSYLTRANSFERASE-I.			
FT	POTENTIAL.			
FT	1			
FT	39			
FT	1590			
FT	175955 NM;			
FT	C3C83A57CF3C280E CRC64;			
FT	39			
FT	1590			
FT	GLUCOSYLTRANSFERASE-I.			
FT	POTENTIAL.			
FT	1			
FT	39			
FT	1590			
FT	175955 NM;			
FT	C3C83A57CF3C280E CRC64;			
FT	39			
FT	1590			
FT	GLUCOSYLTRANSFERASE-I.			
FT	POTENTIAL.			
FT	1			
FT	39			
FT	1590			
FT	175955 NM;			
FT	C3C83A57CF3C280E CRC64;			
FT	39			
FT	1590			
FT	GLUCOSYLTRANSFERASE-I.			
FT	POTENTIAL.			
FT	1			
FT	39			
FT	1590			
FT	175955 NM;			
FT	C3C83A57CF3C280E CRC64;			
FT	39			
FT	1590			
FT	GLUCOSYLTRANSFERASE-I.			
FT	POTENTIAL.			
FT	1			
FT	39			
FT	1590			
FT	175955 NM;			
FT	C3C83A57CF3C280E CRC64;			
FT	39			
FT	1590			
FT	GLUCOSYLTRANSFERASE-I.			
FT	POTENTIAL.			
FT	1			
FT	39			
FT	1590			
FT	175955 NM;			
FT	C3C83A57CF3C280E CRC64;			
FT	39			
FT	1590			
FT	GLUCOSYLTRANSFERASE-I.			
FT	POTENTIAL.			
FT	1			
FT	39			
FT	1590			
FT	175955 NM;			
FT	C3C83A57CF3C280E CRC64;			
FT	39			
FT	1590			
FT	GLUCOSYLTRANSFERASE-I.			
FT	POTENTIAL.			
FT	1			
FT	39			
FT	1590			
FT	175955 NM;			
FT	C3C83A57CF3C280E CRC64;			
FT	39			
FT	1590			
FT	GLUCOSYLTRANSFERASE-I.			
FT	POTENTIAL.			
FT	1			
FT	39			
FT	1590			
FT	175955 NM;			
FT	C3C83A57CF3C280E CRC64;			
FT	39			
FT	1590			
FT	GLUCOSYLTRANSFERASE-I.			
FT	POTENTIAL.			
FT	1			
FT	39			
FT	1590			
FT	175955 NM;			
FT	C3C83A57CF3C280E CRC64;			
FT	39			
FT	1590			
FT	GLUCOSYLTRANSFERASE-I.			
FT	POTENTIAL.			
FT	1			
FT	39			
FT	1590			
FT	175955 NM;			
FT	C3C83A57CF3C280E CRC64;			
FT	39			
FT	1590			
FT	GLUCOSYLTRANSFERASE-I.			
FT	POTENTIAL.			
FT	1			
FT	39			
FT	1590			
FT	175955 NM;			
FT	C3C83A57CF3C280E CRC64;			
FT	39			
FT	1590			
FT	GLUCOSYLTRANSFERASE-I.			
FT	POTENTIAL.			
FT	1			
FT	39			
FT	1590			
FT	175955 NM;			
FT	C3C83A57CF3C280E CRC64;			
FT	39			
FT	1590			
FT	GLUCOSYLTRANSFERASE-I.			
FT	POTENTIAL.			
FT	1			
FT	39			
FT	1590			
FT	175955 NM;			
FT	C3C83A57CF3C280E CRC64;			
FT	39			
FT	1590			
FT	GLUCOSYLTRANSFERASE-I.			
FT	POTENTIAL.			
FT	1			
FT	39			
FT	1590			
FT	175955 NM;			
FT	C3C83A57CF3C280E CRC64;			
FT	39			
FT	1590			
FT	GLUCOSYLTRANSFERASE-I.			
FT	POTENTIAL.			
FT	1			
FT	39			
FT	1590			
FT	175955 NM;			
FT	C3C83A57CF3C280E CRC64;			
FT	39			
FT	1590			
FT	GLUCOSYLTRANSFERASE-I.			
FT	POTENTIAL.			
FT	1			
FT	39			
FT	1590			
FT	175955 NM;			
FT	C3C83A57CF3C280E CRC64;			
FT	39			
FT	1590			
FT	GLUCOSYLTRANSFERASE-I.			
FT	POTENTIAL.			
FT	1			
FT	39			
FT	1590			
FT	175955 NM;			
FT	C3C83A57CF3C280E CRC64;			
FT	39			
FT	1590			
FT	GLUCOSYLTRANSFERASE-I.			
FT	POTENTIAL.			
FT	1			
FT	39			
FT	1590			
FT	175955 NM;			
FT	C3C83A57CF3C280E CRC64;			
FT	39			
FT	1590			
FT	GLUCOSYLTRANSFERASE-I.			
FT	POTENTIAL.			
FT	1			
FT	39			
FT	1590			
FT	175955 NM;			
FT	C3C83A57CF3C280E CRC64;			
FT	39			
FT	1590			
FT	GLUCOSYLTRANSFERASE-I.			
FT	POTENTIAL.			
FT	1			
FT	39			
FT	1590			
FT	175955 NM;			
FT	C3C83A57CF3C280E CRC64;			
FT	39			
FT	1590			
FT	GLUCOSYLTRANSFERASE-I.			
FT	POTENTIAL.			
FT	1			
FT	39			
FT	1590			
FT	175955 NM;			
FT	C3C83A57CF3C280E CRC64;			
FT	39			
FT	1590			
FT	GLUCOSYLTRANSFERASE-I.			
FT	POTENTIAL.			
FT	1			
FT	39			
FT	1590			
FT	175955 NM;			
FT	C3C83A57CF3C280E CRC64;			
FT	39			
FT	1590			
FT	GLUCOSYLTRANSFERASE-I.			
FT	POTENTIAL.			
FT	1			
FT	39			
FT	1590			
FT	175955 NM;			
FT	C3C83A57CF3C280E CRC64;			
FT	39			
FT	1590			
FT	GLUCOSYLTRANSFERASE-I.			
FT	POTENTIAL.			
FT	1			
FT	39			
FT	1590			
FT	175955 NM;			
FT	C3C83A57CF3C280E CRC64;			
FT	39			
FT	1590			
FT	GLUCOSYLTRANSFERASE-I.			
FT	POTENTIAL.			
FT	1			
FT	39			
FT	1590			
FT	175955 NM;			
FT	C3C83A57CF3C280E CRC64;			
FT	39			
FT	1590			
FT	GLUCOSYLTRANSFERASE-I.			
FT	POTENTIAL.			
FT	1			
FT	39			
FT	1590			
FT	175955 NM;			
FT	C3C83A57CF3C280E CRC64;			
FT	39			
FT	1590			
FT	GLUCOSYLTRANSFERASE-I.			
FT	POTENTIAL.			
FT	1			
FT	39			
FT	1590			
FT	175955 NM;			
FT	C3C83A57CF3C280E CRC64;			
FT	39			
FT	1590			
FT	GLUCOSYLTRANSFERASE-I.			
FT	POTENTIAL.			
FT	1			
FT	39			
FT	1590			
FT	175955 NM;			
FT	C3C83A57CF3C280E CRC64;			
FT	39			
FT	1590			
FT	GLUCOSYLTRANSFERASE-I.			
FT	POTENTIAL.			
FT	1			
FT	39			
FT	1590			
FT	175955 NM;			
FT	C3C83A57CF3C280E CRC64;			
FT	39			
FT	1590			
FT	GLUCOSYLTRANSFERASE-I.			

RESULT 2

Q5263 PRELIMINARY; PRT; 1590 AA.
 ID Q5263
 AC Q5263
 DT 01-NOV-1996 (TREMBlrel. 01, Created)
 DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)
 DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
 DE GTF-1.
 GN GLUCOSYLTRANSFERASE.
 OS Streptococcus sobrinus.
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 RX NCBI_TaxID=1310;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 33478;
 RA Sato S.;
 RT "DNA and amino-acid sequences of water-insoluble-glucan synthetase produced from Streptococcus sobrinus ATCC 33478."
 RL Am. Kagoshima Univ. Dental School 16:23-29(1996).
 DR EMBL; D63570; BAA09792.1; -
 DR InterPro; IPR002479; CW_binding.
 DR Pfam; PF01473; Glyco_hydro_70.
 DR Pfam; PF02324; Glyco_hydro_70; 1.
 DR Pfam; PF02324; Glyco_hydro_70; 1.
 SQ SEQUENCE 1590 AA; 176058 MW; 9DF7A3FC6B4FD43 CRC64;

Query Match 89.1%; Score 98; DB 2; Length 1590;
 Best Local Similarity 86.4%; Pred. No. 8.1e-07;
 Matches 19; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 VPSYFTRAHDSVQDLIRNII 22
 |||||
 DB 548 VPSYFTRAHDSVQDLIRNII 569

RESULT 3

Q48756 PRELIMINARY; PRT; 1290 AA.
 ID Q48756
 AC Q48756
 DT 01-NOV-1996 (TREMBlrel. 01, Created)
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
 DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
 DE Dextranucrase.
 OS Leuconostoc mesenteroides.
 OC Bacteria; Firmicutes; Lactobacillales; Leuconostoc.
 RX NCBI_TaxID=1245;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN=NRRL B-1299;
 RA MEDLINE=9713666; PubMed=8982063;
 RA Monchois V., Willemot R.M., Renaud-Simeon M., Groux C., Monsan P.;
 RT "Cloning and sequencing of a gene coding for a novel dextranucrase from Leuconostoc mesenteroides NRRL B-1299 synthesizing only alpha (1-6) and alpha (1-3) linkages."
 RT Gene 182:23-32(1996).
 RL EMBL; U38181; AAB40875.1; -
 DR InterPro; IPR002479; CW_binding.
 DR InterPro; IPR003318; Glyco_hydro_70.
 DR Pfam; PF01473; CW_binding_1; 11.
 DR Pfam; PF02324; Glyco_hydro_70; 1.
 SQ SEQUENCE 1290 AA; 145590 MW; 3555C2E96B749FMA CRC64;

Query Match 75.5%; Score 83; DB 2; Length 1290;
 Best Local Similarity 76.2%; Pred. No. 0.00019;
 Matches 16; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 2 PYSYFTRAHDSVQDLIRNII 22
 |||||
 DB 388 PYSYFTRAHDSVQDLIRNII 408

RESULT 4

Q9L466 PRELIMINARY; PRT; 1477 AA.
 ID Q9L466
 AC Q9L466
 DT 01-OCT-2000 (TREMBlrel. 15, Created)
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
 DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
 DE Dextranucrase (EC 2.4.1.5).
 GN DSRC.
 OS Leuconostoc mesenteroides.
 OC Bacteria; Firmicutes; Lactobacillales; Leuconostoc.
 RX NCBI_TaxID=1245;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN=NRRL B-1355;
 RA Arguello-Morales M.A., Renaud-Simeon M., Pizzut S., Sarcabal P.,
 RA Willemot R.M., Monsan P.;
 RT "Sequence analysis of the gene encoding alternansucrase, a sucrose glucosyltransferase from Leuconostoc mesenteroides NRRL B-1355."
 RT Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
 RL EMBL; A0250172; CAB76565.1; -
 DR InterPro; IPR002479; CW_binding.
 DR InterPro; IPR003318; Glyco_hydro_70.
 DR Pfam; PF01473; CW_binding_1; 14.
 DR Pfam; PF02324; Glyco_hydro_70; 1.
 DR Pfam; PF02324; Glyco_hydro_70; 1.
 KW Glycosyltransferase; Transferase.
 SQ SEQUENCE 1477 AA; 164887 MW; E6F5710DEDFCB831 CRC64;

Query Match 75.5%; Score 83; DB 2; Length 1477;
 Best Local Similarity 68.2%; Pred. No. 0.0002; 3; Indels 0; Gaps 0;
 Matches 15; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 VPSYFTRAHDSVQDLIRNII 22
 |||||
 DB 603 IPYSFVRAHDSVQDLIRNII 624

RESULT 5

Q9EZH5 PRELIMINARY; PRT; 1508 AA.
 ID Q9EZH5
 AC Q9EZH5
 DT 01-MAR-2001 (TREMBlrel. 16, Created)
 DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
 DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
 DE Dextranucrase Dsb742.
 GN DSRB742.
 OS Leuconostoc mesenteroides.
 OC Bacteria; Firmicutes; Lactobacillales; Leuconostoc.
 RX NCBI_TaxID=1245;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN=B-742CB;
 RA Kim H.-S., Kim D., Ryu H.-J., Robyt J.F.;
 RA "Leuconostoc mesenteroides B-742CB, a dextranucrase gene."
 RT Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
 RL EMBL; AF294469; AAG38021.1; -
 DR InterPro; IPR002479; CW_binding.
 DR InterPro; IPR003318; Glyco_hydro_70.
 DR Pfam; PF01473; CW_binding_1; 14.
 DR Pfam; PF02324; Glyco_hydro_70; 1.
 DR Pfam; PF02324; Glyco_hydro_70; 1.
 SQ SEQUENCE 1508 AA; 168542 MW; E2FCFA0F87A4F3A CRC64;

Query Match 75.5%; Score 83; DB 2; Length 1508;
 Best Local Similarity 68.2%; Pred. No. 0.00021;
 Matches 15; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 VPSYFTRAHDSVQDLIRNII 22
 |||||
 DB 634 IPYSFVRAHDSVQDLIRNII 655

RESULT 6

Q52224

ID 052224 PRELIMINARY; PRT: 1508 AA.
 AC 052224;
 DT 01-JUN-1998 (TREMBLrel. 06, Created)
 DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)
 DT 01-OCT-2002 (TREMBLrel. 22, Last annotation update)
 DE Glucosyltransferase (EC 2.4.1.5).
 GN DSRB.
 OS Leuconostoc mesenteroides.
 OC Bacteria; Firmicutes; Lactobacillales; Leuconostoc.
 NCBI_TaxID=1245;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=NRRL B-1299;
 RA Monchois V., Remaud-Simeon M., Monsan P., Willemot R.M.;
 RT "Cloning and sequencing of a gene coding for an extracellular
 dextransucrase (DSRB) from Leuconostoc mesenteroides NRRL B-1299
 synthesizing only a (1-6) glucan."
 RL EMBL; AF030129; AAB95453.1; -
 DR InterPro: IPR002479; Glyco_hydro_70.
 DR InterPro: IPR003318; Glyco_hydro_14.
 DR Pfam; PF01473; CM_binding_1; 1.
 DR Pfam; PF02324; Glyco_hydro_70; 1.
 KM Glycosyltransferase; Transferase.
 SQ SEQUENCE 1508 AA; 168511 MW; E70CECB57A70D1F0 CRC64;

Query Match 75.5%; Score 83; DB 2; Length 1508;
 Best Local Similarity 68.2%; Pred. No. 0.00021;
 Matches 15; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 VPSYSFIRAHDSVQDLIRNII 22
 Db 634 IPNYSFVRADSEVQTVIAQII 655

RESULT 7
 ID 092AR4 PRELIMINARY; PRT: 1527 AA.
 AC 092AR4;
 DT 01-MAY-1999 (TREMBLrel. 10, Created)
 DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)
 DT 01-OCT-2002 (TREMBLrel. 22, Last annotation update)
 DE Dextranucrase.
 GN DEX.
 OS Leuconostoc mesenteroides.
 OC Bacteria; Firmicutes; Lactobacillales; Leuconostoc.
 NCBI_TaxID=1245;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=NRRL B-512-F;
 RA Bhanagar R., Singh D.K.S.;
 RT "Cloning and Molecular Characterization of Dextranucrase Gene from
 Leuconostoc mesenteroides NRRL B-512F."
 RL Submitted (DEC-1996) to the EMBL/Genbank/DBJ databases.
 DR EMBL; U81374; AAD10952.1; -
 DR InterPro: IPR002479; CM_binding_70.
 DR InterPro: IPR003318; Glyco_hydro_70.
 DR Pfam; PF01473; CM_binding_1; 16.
 DR Pfam; PF02324; Glyco_hydro_70; 1.
 KM Transferase.
 SQ SEQUENCE 1527 AA; 169709 MW; IDFAFA237C743398 CRC64;

Query Match 74.5%; Score 82; DB 2; Length 1527;
 Best Local Similarity 63.6%; Pred. No. 0.00031;
 Matches 14; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

QY 1 VPSYSFIRAHDSVQDLIRNII 22
 Db 652 IPNYSFVRADSEVQTVIAQIV 673
 RESULT 8
 Q8KREI PRELIMINARY; PRT: 1527 AA.
 ID Q8KREI

AC Q8KREI;
 DT 01-OCT-2002 (TREMBLrel. 22, Created)
 DT 01-OCT-2002 (TREMBLrel. 22, Last sequence update)
 DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
 DE Dextranucrase DsrD (EC 2.4.1.5).
 GN DSRD.
 OS Leuconostoc mesenteroides.
 OC Bacteria; Firmicutes; Lactobacillales; Leuconostoc.
 NCBI_TaxID=1245;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Neubauer H., Bauche A., Mollet B.;
 RT "Isolation and characterization of the dextransucrase DsrD of
 Leuconostoc mesenteroides Lcc4."
 RL Submitted (JAN-2001) to the EMBL/Genbank/DBJ databases.
 DR EMBL; AY017384; AAG61158.1; -
 DR InterPro: IPR002479; CM_binding_70.
 DR InterPro: IPR003318; Glyco_hydro_70.
 DR Pfam; PF01473; CM_binding_1; 6.
 DR Pfam; PF02324; Glyco_hydro_70; 1.
 KM Transferase; Glycosyltransferase.
 SQ SEQUENCE 1527 AA; 169835 MW; F9DDDE220BD89668 CRC64;

Query Match 74.5%; Score 82; DB 2; Length 1527;
 Best Local Similarity 63.6%; Pred. No. 0.00031;
 Matches 14; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

QY 1 VPSYSFIRAHDSVQDLIRNII 22
 Db 652 IPNYSFVRADSEVQTVIAQIV 673

RESULT 9
 ID Q8KZL5 PRELIMINARY; PRT: 1554 AA.
 AC Q8KZL5;
 DT 01-OCT-2002 (TREMBLrel. 22, Created)
 DT 01-OCT-2002 (TREMBLrel. 22, Last sequence update)
 DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
 DE Glucosyltransferase.
 GN GTFU.
 OS Streptococcus sobrinus.
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 NCBI_TaxID=1310;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21958684; Pubmed=11960691;
 RA Hanada N., Fukushima K., Nomura Y., Sempuku H., Hayakawa M.,
 RA Mukasa H., Shiroza T., Abiko Y.;
 RT "Cloning and nucleotide sequence analysis of the Streptococcus
 sobrinus gtfu gene that produces a highly branched water-soluble
 glucan."
 RL Biochim. Biophys. Acta 1570:75-79 (2002).
 DR EMBL; AB089438; BAC07265.1; -
 DR InterPro: IPR002479; CM_binding_70.
 DR InterPro: IPR003318; Glyco_hydro_70.
 DR Pfam; PF01473; CM_binding_1; 6.
 DR Pfam; PF02324; Glyco_hydro_70; 1.
 KM Transferase.
 SQ SEQUENCE 1554 AA; 171676 MW; 6981BCC1DAE24A73 CRC64;

Query Match 74.5%; Score 82; DB 2; Length 1554;
 Best Local Similarity 63.6%; Pred. No. 0.00031;
 Matches 14; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

QY 1 VPSYSFIRAHDSVQDLIRNII 22
 Db 557 IPNYSFVRADSEVQTVIAKIV 578
 RESULT 10
 Q9LCJ7

ID 09LCU7; PRELIMINARY; PRT; 1016 AA.
 AC 09LCU7; 01-OCT-2000 (Tremblrel. 15, Created)
 DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
 DT 01-OCT-2002 (Tremblrel. 22, Last annotation update)
 DE Dextranucrase.
 GN DSRT.
 OS Leuconostoc mesenteroides.
 OC Bacteria; Firmicutes; Lactobacillales; Leuconostoc.
 NCBI_TaxID=1245;
 RX NCBI_TaxID=1245;
 RN SEQUENCE FROM N.A.
 RC STRAIN=NRRL-B-512F;
 RX MEDLINE=20169623; PubMed=10705445; Kobayashi M.;
 RA Funane K., Mizuno K., Takahara H., Kobayashi M.;
 RT "gene encoding a dextranucrase-like protein in Leuconostoc
 RT mesenteroides NRRL-B-512F.";
 RL Biosci. Biotechnol. Biochem. 64:29-38(2000).
 DR EMBL; AB020020; BAA90527.1; -.
 DR HSSP; P06278; 1VTS.
 DR InterPro; IPR003318; Glyco_hydro_70.
 DR Pfam; PF02324; Glyco_hydro_70; 1.
 SQ SEQUENCE 1016 AA; 110344 MW; 8896FDEB13CCB47 CRC64;
 QY Query Match 72.7%; Score 80; DB 2; Length 1016;
 Best Local Similarity 71.4%; Pred. No. 0.00041;
 Matches 15; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

2 PSYSEFRAHDESVQDLIRNII 22
 DB 625 PNYAFRAHDESVQTVIAQII 645

RESULT 11
 ID 08G9Q2; PRELIMINARY; PRT; 2835 AA.
 AC 08G9Q2; 01-MAR-2003 (Tremblrel. 23, Created)
 DT 01-MAR-2003 (Tremblrel. 23, Last sequence update)
 DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
 DE Dextranucrase (EC 2.4.1.5) (Fragment).
 GN DSRE.
 OS Leuconostoc mesenteroides.
 OC Bacteria; Firmicutes; Lactobacillales; Leuconostoc.
 NCBI_TaxID=1245;
 RX NCBI_TaxID=1245;
 RN SEQUENCE FROM N.A.
 RC MEDLINE=2231661; PubMed=12270834;
 RA Bozonnet S., Dols-Lafargue M., Fabre E., Pizut S., Renaud-Simeon M.,
 RT "Molecular characterization of DSR-E, an alpha-1,2 linkage
 RT synthesizing dextranucrase with two catalytic domains.";
 RL J. Bacteriol. 184:5753-5761(2002).
 DR EMBL; AJ430204; CAD22883.1; -.
 DR Transferase; Glycosyltransferase.
 KW NON TER 1
 SQ SEQUENCE 2835 AA; 313264 MW; D03262CDD735399D CRC64;
 QY Query Match 70.9%; Score 78; DB 2; Length 2835;
 Best Local Similarity 71.4%; Pred. No. 0.00277;
 Matches 15; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

2 PSYSEFRAHDESVQDLIRNII 22
 DB 629 PNYAFRAHDESVQTVIAQII 649

RESULT 12
 ID 09MXJ5; PRELIMINARY; PRT; 1512 AA.
 AC 09MXJ5; 01-NOV-1999 (Tremblrel. 12, Created)
 DT 01-NOV-1999 (Tremblrel. 12, Last sequence update)
 DT 01-NOV-1999 (Tremblrel. 12, Last sequence update)

DT 01-OCT-2002 (Tremblrel. 22, Last annotation update)
 DE GTF-S.
 GN GTF.
 OS Streptococcus criceti.
 OC Plasmid pAM1.
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 CC Streptococcus.
 NCBI_TaxID=1333;
 RX NCBI_TaxID=1333;
 RN SEQUENCE FROM N.A.
 RC STRAIN=HS-6;
 RA Inoue M., Fukui K., Miyagi A.;
 RT "S. cricetus glucosyltransferase (gifs and gtf) genes.";
 RT Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
 RL EMBL; AB026123; BAA77237.1; -.
 DR InterPro; IPR002479; CW_binding.
 DR InterPro; IPR003318; Glyco_hydro_70.
 DR Pfam; PF01473; CW_binding_1; 14.
 DR Pfam; PF02324; Glyco_hydro_70; 1.
 KW Plasmid.
 SQ SEQUENCE 1512 AA; 167145 MW; 4C03D9C8C601FC14 CRC64;
 QY Query Match 70.0%; Score 77; DB 2; Length 1512;
 Best Local Similarity 71.4%; Pred. No. 0.002;
 Matches 15; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

2 PSYSEFRAHDESVQDLIRNII 22
 DB 560 PSYVFRADHSEVQTVIAQII 580

RESULT 13
 ID 09MXJ4; PRELIMINARY; PRT; 1338 AA.
 AC 09MXJ4; 01-NOV-1999 (Tremblrel. 12, Created)
 DT 01-NOV-1999 (Tremblrel. 12, Last sequence update)
 DT 01-OCT-2002 (Tremblrel. 22, Last annotation update)
 DE GTF-S.
 GN GTF-S.
 OS Streptococcus criceti.
 OS Plasmid pAM1.
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 CC Streptococcus.
 NCBI_TaxID=1333;
 RX NCBI_TaxID=1333;
 RN SEQUENCE FROM N.A.
 RC STRAIN=HS-6;
 RA Inoue M., Fukui K., Miyagi A.;
 RT "S. cricetus glucosyltransferase (gifs and gtf) genes.";
 RT Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
 RL EMBL; AB026123; BAA77237.1; -.
 DR InterPro; IPR002479; CW_binding.
 DR InterPro; IPR003318; Glyco_hydro_70.
 DR Pfam; PF01473; CW_binding_1; 10.
 DR Pfam; PF02324; Glyco_hydro_70; 1.
 KW Plasmid.
 SQ SEQUENCE 1338 AA; 148558 MW; 0A90C8E10E15D99B CRC64;
 QY Query Match 69.1%; Score 76; DB 2; Length 1338;
 Best Local Similarity 72.7%; Pred. No. 0.0025;
 Matches 16; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

1 PSYSEFRAHDESVQDLIRNII 22
 DB 509 PSYVFRADHSEVQTVIAQII 530

RESULT 14
 ID 068542; PRELIMINARY; PRT; 1449 AA.
 AC 068542; 01-AUG-1998 (Tremblrel. 07, Created)
 DT 01-AUG-1998 (Tremblrel. 07, Created)

```

DT 01-AUG-1998 (TRENBLREL. 07, last sequence update)
DT 01-OCT-2002 (TRENBLREL. 22, last annotation update)
DE Glucosyltransferase N (Fragment).
GN GTFN.
OS Streptococcus salivarius.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1304;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=V1477;
RA Jaffe R.I.;
RT "Streptococcus salivarius V1477 gtfN."
RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF049609; AAC05156.1; -.
DR InterPro; IPR002479; CW_binding.
DR InterPro; IPR003318; Glyco_hydro_70.
DR Pfam; PF01473; CW_binding_1; 8.
DR Pfam; PF02324; Glyco_hydro_70; 1.
KW Transferase.
FT NON_TER 1449 1449
SQ SEQUENCE 1449 AA; 159895 MW; 0700FED748471BFB CRC64;

Query Match 64.5%; Score 71; DB 2; Length 1449;
Best Local Similarity 65.0%; Pred. No. 0.018;
Matches 13; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 3 SYSFIRAHDSVODLRNII 22
Db 609 NYAFVRAHDSVOSIIIGI 628

RESULT 15
Q55264 PRELIMINARY; PRT; 1449 AA.
ID Q55264
AC Q55264;
DT 01-NOV-1996 (TRENBLREL. 01, Created)
DT 01-NOV-1996 (TRENBLREL. 01, last sequence update)
DT 01-OCT-2002 (TRENBLREL. 22, last annotation update)
DE Glucosyltransferase precursor.
GN GTFP.
OS Streptococcus salivarius.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1304;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95122197; PubMed=7822030;
RA Simpson C.L., Giffard P.M., Jacques N.A.;
RT "Streptococcus salivarius ATCC 25975 possesses at least two genes
RT coding for primer-independent glucosyltransferases."
RL Infect. Immun. 63:609-621 (1995).
DR EMBL; U35495; AAC41412.1; -.
DR InterPro; IPR002479; CW_binding.
DR InterPro; IPR003318; Glyco_hydro_70.
DR Pfam; PF01473; CW_binding_1; 8.
DR Pfam; PF02324; Glyco_hydro_70; 1.
KW Signal; Transferase.
FT SIGNAL 1 35
FT CHAIN 36 1449 POTENTIAL.
SQ SEQUENCE 1449 AA; 159984 MW; DD62F07306E86A46 CRC64;

Query Match 64.5%; Score 71; DB 2; Length 1449;
Best Local Similarity 65.0%; Pred. No. 0.018;
Matches 13; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 3 SYSFIRAHDSVODLRNII 22
Db 609 NYAFVRAHDSVOSIIIGI 628

```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

CM protein - protein search, using sw model

Run on: November 13, 2003, 09:11:40 ; Search time 14.2844 Seconds
(without alignments)
65.165 Million cell updates/sec

Title: US-09-290-049A-15
Perfect score: 110
Sequence: 1 VPSYSFIRAHDSVQDLIRNII 22

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 328717 segs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/prodata/2/iaa/5A COMB pep:*
2: /cgn2_6/prodata/2/iaa/5B COMB pep:*
3: /cgn2_6/prodata/2/iaa/6A COMB pep:*
4: /cgn2_6/prodata/2/iaa/6B COMB pep:*
5: /cgn2_6/prodata/2/iaa/PCUTS COMB pep:*
6: /cgn2_6/prodata/2/iaa/backfile1 pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	110	100.0	1375	3 US-09-210-361-4	Sequence 4, Appli
2	110	100.0	1375	4 US-09-740-274-4	Sequence 4, Appli
3	99	90.0	1475	3 US-09-007-999-2	Sequence 2, Appli
4	99	90.0	1475	3 US-09-210-361-2	Sequence 2, Appli
5	99	90.0	1475	4 US-09-740-274-2	Sequence 2, Appli
6	82	74.5	523	4 US-09-604-957-5	Sequence 5, Appli
7	76	69.1	535	4 US-09-604-957-7	Sequence 7, Appli
8	76	69.1	1278	4 US-09-604-957-3	Sequence 7, Appli
9	71	64.5	2057	4 US-09-499-203-2	Sequence 2, Appli
10	70	63.6	1577	2 US-08-793-824-2	Sequence 2, Appli
11	68	61.8	545	4 US-09-604-957-4	Sequence 2, Appli
12	68	61.8	1430	3 US-09-008-172-2	Sequence 6, Appli
13	68	61.8	1430	3 US-09-210-361-6	Sequence 6, Appli
14	68	61.8	1430	4 US-09-740-274-6	Sequence 6, Appli
15	66	60.0	584	4 US-09-604-957-6	Sequence 6, Appli
16	50	45.5	565	4 US-09-107-532A-4217	Sequence 4217, Ap
17	47	42.7	639	4 US-09-509-814A-4	Sequence 4, Appli
18	46	41.8	196	2 US-08-684-024-2	Sequence 2, Appli
19	46	41.8	196	2 US-08-684-024-2	Sequence 9, Appli
20	46	41.8	196	3 US-09-145-868-2	Sequence 9, Appli
21	46	41.8	196	3 US-09-145-868-9	Sequence 9, Appli
22	44	40.0	282	4 US-09-172-952-20	Sequence 20, Appli
23	44	40.0	1068	4 US-08-390-874C-11	Sequence 11, Appli
24	44	40.0	1068	4 US-09-265-772-11	Sequence 11, Appli
25	44	40.0	1069	2 US-08-162-081B-37	Sequence 37, Appli
26	44	40.0	1069	2 US-08-780-872-37	Sequence 37, Appli
27	44	40.0	1069	3 US-09-085-957-37	Sequence 37, Appli

28	44	40.0	1080	2 US-08-162-081B-36	Sequence 36, Appli
29	44	40.0	1080	2 US-08-780-872-36	Sequence 36, Appli
30	44	40.0	1080	3 US-09-085-957-36	Sequence 36, Appli
31	43	39.1	75	4 US-09-107-532A-7020	Sequence 7020, Ap
32	43	39.1	347	4 US-09-461-325-277	Sequence 277, Ap
33	42	38.2	211	4 US-09-328-352-4995	Sequence 4995, Ap
34	42	38.2	2627	2 US-08-751-189-3	Sequence 3, Appli
35	42	38.2	2627	2 US-08-060-836-3	Sequence 3, Appli
36	42	38.2	2627	3 US-09-184-445-3	Sequence 3, Appli
37	41.5	37.7	348	4 US-09-252-991A-31249	Sequence 31249, A
38	41.5	37.7	713	2 US-08-849-212-4	Sequence 4, Appli
39	41	37.3	616	4 US-09-268-347-38	Sequence 38, Appli
40	41	37.3	659	4 US-09-268-347-46	Sequence 46, Appli
41	41	37.3	660	4 US-09-268-347-45	Sequence 45, Appli
42	40	36.4	309	4 US-09-345-473E-37	Sequence 37, Appli
43	40	36.4	472	2 US-08-272-255-10	Sequence 10, Appli
44	40	36.4	472	3 US-08-964-268-4	Sequence 4, Appli
45	40	36.4	472	4 US-09-105-254-4	Sequence 4, Appli

ALIGNMENTS

```

RESULT 1
US-09-210-361-4
; Sequence 4, Application US/09210361
; Patent No. 6284479
; GENERAL INFORMATION:
; APPLICANT: Nichols, Scott E.
; TITLE OF INVENTION: Substitutes for Modified Starches and
; FILE REFERENCE: 0357CR
; CURRENT APPLICATION NUMBER: US/09/210,361
; EARLIER FILING DATE: 1998-12-11
; EARLIER APPLICATION NUMBER: 09/007,999
; EARLIER FILING DATE: 1998-01-16
; EARLIER APPLICATION NUMBER: 08/478,704
; EARLIER FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 09/009,620
; EARLIER FILING DATE: 1998-01-20
; EARLIER APPLICATION NUMBER: 08/485,243
; EARLIER FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 09/008,172
; EARLIER FILING DATE: 1998-01-16
; EARLIER APPLICATION NUMBER: 08/482,711
; EARLIER FILING DATE: 1995-06-07
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 1375
; TYPE: PRT
; ORGANISM: streptococcus mutans
US-09-210-361-4

Query Match          100.0%; Score 110; DB 3; Length 1375;
Best local Similarity 100.0%; Pred. No. 3.4e-10;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
      1 VPSYSFIRAHDSVQDLIRNII 22
      |||||
Db      578 VPSYSFIRAHDSVQDLIRNII 599

RESULT 2
US-09-740-274-4
; Sequence 4, Application US/09740274
; Patent No. 6465203
; GENERAL INFORMATION:
; APPLICANT: Nichols, Scott E.
; TITLE OF INVENTION: Glucan-containing Compositions and Paper
; FILE REFERENCE: 0357CRD
; CURRENT APPLICATION NUMBER: US/09/740,274
; CURRENT FILING DATE: 2000-12-19

```

```

; PRIOR APPLICATION NUMBER: 09/210,361
; PRIOR FILING DATE: 1998-12-11
; PRIOR APPLICATION NUMBER: 09/007,999
; PRIOR FILING DATE: 1998-01-16
; PRIOR APPLICATION NUMBER: 08/478,704
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: 09/009,620
; PRIOR FILING DATE: 1998-01-20
; PRIOR APPLICATION NUMBER: 08/485,243
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: 09/008,172
; PRIOR FILING DATE: 1998-01-16
; PRIOR APPLICATION NUMBER: 08/482,711
; PRIOR FILING DATE: 1995-06-07
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 1375
; TYPE: PRF
; ORGANISM: Streptococcus mutans
US-09-740-274-4
```

```

Query Match          100.0%; Score 110; DB 4; Length 1375;
Best Local Similarity 100.0%; Pred. No. 3.4e-10;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 VPSYSFIRAHDSVQDLIRNII 22
      |||||
Db      578 VPSYSFIRAHDSVQDLIRNII 599
```

```

RESULT 3
US-09-007-999-2
; Sequence 2, Application US/09007999
; Patent No. 6087559
; GENERAL INFORMATION:
; APPLICANT: Nichols, Scott E.
; TITLE OF INVENTION: Substitutes for Modified Starch and
; FILE REFERENCE: 0356D
; CURRENT APPLICATION NUMBER: US/09/007,999
; CURRENT FILING DATE: 1998-01-16
; EARLIER APPLICATION NUMBER: 08/478,704
; EARLIER FILING DATE: 1995-06-07
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 1475
; TYPE: PRF
; ORGANISM: Streptococcus mutans
US-09-007-999-2
```

```

Query Match          90.0%; Score 99; DB 3; Length 1475;
Best Local Similarity 90.9%; Pred. No. 2.9e-08;
Matches 20; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
```

```
QY      1 VPSYSFIRAHDSVQDLIRNII 22
      |||||
Db      552 VPSYSFIRAHDSVQDLIRNII 573
```

```

RESULT 4
US-09-210-361-2
; Sequence 2, Application US/09210361
; Patent No. 6284479
; GENERAL INFORMATION:
; APPLICANT: Nichols, Scott E.
; TITLE OF INVENTION: Substitutes for Modified Starches and
; FILE REFERENCE: 0357CR
; CURRENT APPLICATION NUMBER: US/09/210,361
; CURRENT FILING DATE: 1998-12-11
; EARLIER APPLICATION NUMBER: 09/007,999
```

```

; EARLIER FILING DATE: 1998-01-16
; EARLIER APPLICATION NUMBER: 08/478,704
; EARLIER FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 09/009,620
; EARLIER FILING DATE: 1998-01-20
; EARLIER APPLICATION NUMBER: 08/485,243
; EARLIER FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 09/008,172
; EARLIER FILING DATE: 1998-01-16
; EARLIER APPLICATION NUMBER: 08/482,711
; EARLIER FILING DATE: 1995-06-07
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 1475
; TYPE: PRF
; ORGANISM: Streptococcus mutans
US-09-210-361-2
```

```

Query Match          90.0%; Score 99; DB 4; Length 1475;
Best Local Similarity 90.9%; Pred. No. 2.9e-08;
Matches 20; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
```

```
QY      1 VPSYSFIRAHDSVQDLIRNII 22
      |||||
Db      552 VPSYSFIRAHDSVQDLIRNII 573
```

```

RESULT 5
US-09-740-274-2
; Sequence 2, Application US/09740274
; Patent No. 6465203
; GENERAL INFORMATION:
; APPLICANT: Nichols, Scott E.
; TITLE OF INVENTION: Glucan-containing Compositions and Paper
; FILE REFERENCE: 0357CRD
; CURRENT APPLICATION NUMBER: US/09/740,274
; CURRENT FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: 09/210,361
; PRIOR FILING DATE: 1998-12-11
; PRIOR APPLICATION NUMBER: 09/007,999
; PRIOR FILING DATE: 1998-01-16
; PRIOR APPLICATION NUMBER: 08/478,704
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: 09/009,620
; PRIOR FILING DATE: 1998-01-20
; PRIOR APPLICATION NUMBER: 08/485,243
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: 09/008,172
; PRIOR FILING DATE: 1998-01-16
; PRIOR APPLICATION NUMBER: 08/482,711
; PRIOR FILING DATE: 1995-06-07
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 1475
; TYPE: PRF
; ORGANISM: Streptococcus mutans
US-09-740-274-2
```

```

Query Match          90.0%; Score 99; DB 4; Length 1475;
Best Local Similarity 90.9%; Pred. No. 2.9e-08;
Matches 20; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
```

```
QY      1 VPSYSFIRAHDSVQDLIRNII 22
      |||||
Db      552 VPSYSFIRAHDSVQDLIRNII 573
```

```

RESULT 6
US-09-604-957-5
; Sequence 5, Application US/09604957
; Patent No. 6486314
```

```

; GENERAL INFORMATION:
; APPLICANT: VAN GEEL-SCHUTTEN, GERRITDINA HENDRIKA
; APPLICANT: DIKHUIZEN, LUBBERT
; APPLICANT: RAHAOUI, HAKIM
; APPLICANT: LEER, ROBERT-JAN
; TITLE OF INVENTION: NOVEL GLUCOSYLTRANSFERASE AND NOVEL GLUCAN
; FILE REFERENCE: BO 43388
; CURRENT APPLICATION NUMBER: US/09/604,957
; CURRENT FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: 00201871.1
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 523
; TYPE: PRT
; ORGANISM: Leuconostoc mesenteroides
US-09-604-957-5

Query Match          74.5%; Score 82; DB 4; Length 523;
Best Local Similarity 63.6%; Pred. No. 6.8e-06;
Matches 14; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

QY      1 VPSYSFRAHDSVQDLIRNII 22
DB      146 IPNYSFRAHDSVQTLTAQIV 167

RESULT 7
US-09-604-957-7
; Sequence 7, Application US/09604957
; Patent No. 6486314
; GENERAL INFORMATION:
; APPLICANT: VAN GEEL-SCHUTTEN, GERRITDINA HENDRIKA
; APPLICANT: DIKHUIZEN, LUBBERT
; APPLICANT: RAHAOUI, HAKIM
; APPLICANT: LEER, ROBERT-JAN
; TITLE OF INVENTION: NOVEL GLUCOSYLTRANSFERASE AND NOVEL GLUCAN
; FILE REFERENCE: BO 43388
; CURRENT APPLICATION NUMBER: US/09/604,957
; CURRENT FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: 00201871.1
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 7
; LENGTH: 535
; TYPE: PRT
; ORGANISM: Lactobacillus reuteri
US-09-604-957-7

Query Match          69.1%; Score 76; DB 4; Length 535;
Best Local Similarity 59.1%; Pred. No. 7.4e-05;
Matches 13; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY      1 VPSYSFRAHDSVQDLIRNII 22
DB      144 IPNYSFRAHDNSQDQIQNAI 165

RESULT 8
US-09-604-957-3
; Sequence 3, Application US/09604957
; Patent No. 6486314
; GENERAL INFORMATION:
; APPLICANT: VAN GEEL-SCHUTTEN, GERRITDINA HENDRIKA
; APPLICANT: DIKHUIZEN, LUBBERT
; APPLICANT: RAHAOUI, HAKIM
; APPLICANT: LEER, ROBERT-JAN
; TITLE OF INVENTION: NOVEL GLUCOSYLTRANSFERASE AND NOVEL GLUCAN
; FILE REFERENCE: BO 43388
; CURRENT APPLICATION NUMBER: US/09/604,957
; CURRENT FILING DATE: 2000-06-28
```

```

; PRIOR APPLICATION NUMBER: 00201871.1
; PRIOR FILING DATE: 2000-05-25
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 1278
; TYPE: PRT
; ORGANISM: Lactobacillus reuteri
US-09-604-957-3

Query Match          69.1%; Score 76; DB 4; Length 1278;
Best Local Similarity 59.1%; Pred. No. 0.00021;
Matches 13; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY      1 VPSYSFRAHDSVQDLIRNII 22
DB      620 IPNYSFRAHDNSQDQIQNAI 641

RESULT 9
US-09-499-203-2
; Sequence 2, Application US/09499203
; Patent No. 6570065
; GENERAL INFORMATION:
; APPLICANT: KOSMANN, Jens
; APPLICANT: WELSH, Thomas
; APPLICANT: QUANZ, Martin
; APPLICANT: KNUTH, Karola
; TITLE OF INVENTION: Nucleic Acid Molecules Encoding Alternansucrase
; FILE REFERENCE: 147-196P
; CURRENT APPLICATION NUMBER: US/09/499,203
; CURRENT FILING DATE: 2000-02-08
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 2057
; TYPE: PRT
; ORGANISM: Leuconostoc mesenteroides
US-09-499-203-2

Query Match          64.5%; Score 71; DB 4; Length 2057;
Best Local Similarity 63.2%; Pred. No. 0.0027;
Matches 12; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY      1 VPSYSFRAHDSVQDLIR 19
DB      757 IPNYSFRAHDYDAQDPIR 775

RESULT 10
US-08-793-824-2
; Sequence 2, Application US/08793824
; Patent No. 5981838
; GENERAL INFORMATION:
; APPLICANT: Simpson, Christine Lynn
; APPLICANT: Giffard, Philip Morrison
; APPLICANT: Jacques, Nicholas Anthony
; TITLE OF INVENTION: Genetic Manipulation of Plants to
; TITLE OF INVENTION: Increase Stored Carbohydrates
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Griffith Hack & Co
; STREET: Level 8, 168 Walker Street
; CITY: No. 5981838th Sydney
; STATE: New South Wales
; COUNTRY: Australia
; ZIP: 2060
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
```

APPLICATION NUMBER: US/08/793,824
FILING DATE:
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: AU PM7643
FILING DATE: 24-AUG-1994
TELECOMMUNICATION INFORMATION:
TELEPHONE: 61 2 9957 5944
TELEFAX: 61 2 957 6288
TELEX: 26547
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1577 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: not relevant
MOLECULE TYPE: protein
ORIGINAL SOURCE:
ORGANISM: Streptococcus salivarius
US-08-793-824-2

Query Match 63.6%; Score 70; DB 2; Length 1577;
Best Local Similarity 65.0%; Pred. No. 0.0029;
Matches 13; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 3 SYSFIRAHSEVQDILRNII 22
: ||||| :
DB 661 NYIFIRAHSEVQAVLANII 660

RESULT 11
US-09-604-957-4
Sequence 4, Application US/09604957
Patent No. 6486314
GENERAL INFORMATION:
APPLICANT: VAN GEEL-SCHUTTEN, GERRITDIJMA HENDRIKA
APPLICANT: DIJKHUIZEN, LOBBERT
APPLICANT: RAHAOUI, HAKIM
APPLICANT: LEER, ROBERT-JAN
TITLE OF INVENTION: NOVEL GLUCOSYLTRANSFERASE AND NOVEL GLUCAN
FILE REFERENCE: BO 43388
CURRENT APPLICATION NUMBER: US/09/604,957
CURRENT FILING DATE: 2000-06-28
PRIOR APPLICATION NUMBER: 00201871.1
PRIOR FILING DATE: 2000-05-25
NUMBER OF SEQ ID NOS: 17
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 4
LENGTH: 545
TYPE: PRT
ORGANISM: Streptococcus mutans
US-09-604-957-4

Query Match 61.8%; Score 68; DB 4; Length 545;
Best Local Similarity 70.0%; Pred. No. 0.0018;
Matches 14; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 3 SYSFIRAHSEVQDILRNII 22
: ||||| :
DB 156 NYIFIRAHSEVQTVIAKII 175

RESULT 12
US-09-008-172-2
Sequence 2, Application US/09008172
Patent No. 6127602
GENERAL INFORMATION:
APPLICANT: NICHOLS, SCOTT E.
TITLE OF INVENTION: Substitutes for Modified Starch and
TITLE OF INVENTION: latexes in Paper Manufacture
FILE REFERENCE: 0358D
CURRENT APPLICATION NUMBER: US/09/008,172
PRIOR FILING DATE: 1998-01-16

EARLIER APPLICATION NUMBER: 08/482,711
EARLIER FILING DATE: 1995-06-07
NUMBER OF SEQ ID NOS: 2
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 2
LENGTH: 1430
TYPE: PRT
ORGANISM: Streptococcus mutans
US-09-008-172-2

Query Match 61.8%; Score 68; DB 3; Length 1430;
Best Local Similarity 70.0%; Pred. No. 0.0057;
Matches 14; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 3 SYSFIRAHSEVQDILRNII 22
: ||||| :
DB 576 NYIFIRAHSEVQTVIAKII 595

RESULT 13
US-09-210-361-6
Sequence 6, Application US/09210361
Patent No. 6284479
GENERAL INFORMATION:
APPLICANT: NICHOLS, SCOTT E.
TITLE OF INVENTION: Substitutes for Modified Starches and
TITLE OF INVENTION: latexes in Paper Manufacture
FILE REFERENCE: 0357CR
CURRENT APPLICATION NUMBER: US/09/210,361
CURRENT FILING DATE: 1998-12-11
EARLIER APPLICATION NUMBER: 09/007,999
EARLIER FILING DATE: 1998-01-16
EARLIER APPLICATION NUMBER: 08/478,704
EARLIER FILING DATE: 1995-06-07
EARLIER APPLICATION NUMBER: 09/009,620
EARLIER FILING DATE: 1998-01-20
EARLIER APPLICATION NUMBER: 08/485,243
EARLIER FILING DATE: 1995-06-07
EARLIER APPLICATION NUMBER: 09/008,172
EARLIER FILING DATE: 1998-01-16
EARLIER APPLICATION NUMBER: 08/482,711
EARLIER FILING DATE: 1995-06-07
NUMBER OF SEQ ID NOS: 6
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 6
LENGTH: 1430
TYPE: PRT
ORGANISM: streptococcus mutans
US-09-210-361-6

Query Match 61.8%; Score 68; DB 3; Length 1430;
Best Local Similarity 70.0%; Pred. No. 0.0057;
Matches 14; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 3 SYSFIRAHSEVQDILRNII 22
: ||||| :
DB 576 NYIFIRAHSEVQTVIAKII 595

RESULT 14
US-09-740-274-6
Sequence 6, Application US/09740274
Patent No. 6465203
GENERAL INFORMATION:
APPLICANT: NICHOLS, SCOTT E.
TITLE OF INVENTION: Glucan-containing Compositions and Paper
FILE REFERENCE: 0357CRD
CURRENT APPLICATION NUMBER: US/09/740,274
CURRENT FILING DATE: 2000-12-19
PRIOR APPLICATION NUMBER: 09/210,361
PRIOR FILING DATE: 1998-12-11
PRIOR APPLICATION NUMBER: 09/007,999
PRIOR FILING DATE: 1998-01-16

; PRIOR APPLICATION NUMBER: 08/478,704
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: 09/009,620
; PRIOR FILING DATE: 1998-01-20
; PRIOR APPLICATION NUMBER: 08/485,243
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: 09/008,172
; PRIOR FILING DATE: 1998-01-16
; PRIOR APPLICATION NUMBER: 08/482,711
; PRIOR FILING DATE: 1995-06-07
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 6
; LENGTH: 1430
; TYPE: PR1
; ORGANISM: streptococcus mutans
US-09-740-274-6

Query Match 61.8%; Score 68; DB 4; Length 1430;
Best Local Similarity 70.0%; Pred. No. 0.0057;
Matches 14; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 3 SYSFIRAHSEVQDLIRNII 22
Db 576 NTFIRAHSEVQTVIAKII 595

RESULT 15
US-09-604-957-6
; Sequence 6, Application US/09604957
; Patent No. 6486314
; GENERAL INFORMATION:
; APPLICANT: VAN GEEL-SCHUTTEN, GERRITDINA HENDRIKA
; APPLICANT: DIJKHUIZEN, LUBBERT
; APPLICANT: RAHAOUT, HAKIM
; APPLICANT: LEER, ROBERT-JAN
; TITLE OF INVENTION: NOVEL GLUCOSYLTRANSFERASE AND NOVEL GLUCAN
; FILE REFERENCE: BO 43388
; CURRENT APPLICATION NUMBER: US/09/604,957
; CURRENT FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: 00201871.1
; PRIOR FILING DATE: 2000-05-25
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 584
; TYPE: PR1
; ORGANISM: Leuconostoc mesenteroides
US-09-604-957-6

Query Match 60.0%; Score 66; DB 4; Length 584;
Best Local Similarity 61.1%; Pred. No. 0.0043;
Matches 11; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 VPSYFIRAHSEVQDLI 18
Db 167 IPNYSFVRAHDYDAQDP1 184

Search Completed: November 13, 2003, 09:48:01
Job time : 14.2844 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 13, 2003, 09:45:40 ; Search time 28.7773 Seconds

(without alignments)
139.565 Million cell updates/sec

Title: US-09-290-049A-15

Perfect score: 110
Sequence: 1 VPSYSTRADSEVQDLIRNII 22

Scoring table:
BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 666188 seqs, 182559486 residues

Total number of hits satisfying chosen parameters: 666188

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications AA:*
1: /cgn2_6/ptodata/2/pubppaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/2/pubppaa/PTCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/2/pubppaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/2/pubppaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/2/pubppaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/2/pubppaa/PTCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/2/pubppaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/2/pubppaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/2/pubppaa/US09_PUBCOMB.pep:*
10: /cgn2_6/ptodata/2/pubppaa/US09C_PUBCOMB.pep:*
11: /cgn2_6/ptodata/2/pubppaa/US09C_NEW_PUB.pep:*
12: /cgn2_6/ptodata/2/pubppaa/US09C_PUBCOMB.pep:*
13: /cgn2_6/ptodata/2/pubppaa/US10_PUBCOMB.pep:*
14: /cgn2_6/ptodata/2/pubppaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/2/pubppaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/2/pubppaa/US10C_NEW_PUB.pep:*
17: /cgn2_6/ptodata/2/pubppaa/US60_NEW_PUB.pep:*
18: /cgn2_6/ptodata/2/pubppaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	110.0	1375	9	US-09-740-274-4	Sequence 4, Appl1
2	99	1475	9	US-09-740-274-2	Sequence 2, Appl1
3	82	74.5	522	US-09-995-749A-11	Sequence 11, Appl1
4	76	69.1	535	US-09-995-749A-13	Sequence 13, Appl1
5	76	69.1	1781	US-09-995-749A-2	Sequence 2, Appl1
6	71	64.5	584	US-09-995-749A-12	Sequence 12, Appl1
7	68	61.8	545	US-09-995-749A-10	Sequence 10, Appl1
8	68	61.8	1430	US-09-740-274-6	Sequence 6, Appl1
9	44	40.0	282	US-09-557-796-20	Sequence 20, Appl1
10	44	40.0	1068	US-10-101-235A-8	Sequence 8, Appl1
11	43	39.1	347	US-10-012-542-277	Sequence 277, Appl1
12	43	39.1	568	US-10-156-761-11791	Sequence 11791, A
13	43	39.1	2541	US-10-177-293-470	Sequence 470, App
14	41.5	37.7	1213	US-09-815-242-5358	Sequence 5358, Ap
15	41.5	37.7	1217	US-09-815-242-12523	Sequence 12523, A

	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45
	40.5	40.5	40.5	40.5	40.5	40.5	40.5	40.5	40.5	40.5	40.5	40.5	40.5	40.5	40.5	40.5	40.5	40.5	40.5	40.5	40.5	40.5	40.5	40.5	40.5	40.5	40.5	40.5	40.5	40.5
	36.8	36.8	36.8	36.8	36.8	36.8	36.8	36.8	36.8	36.8	36.8	36.8	36.8	36.8	36.8	36.8	36.8	36.8	36.8	36.8	36.8	36.8	36.8	36.8	36.8	36.8	36.8	36.8	36.8	36.8
	296	481	487	43	309	472	489	572	573	574	574	779	443	67	108	188	228	405	557	1383	1383	1383	1389	1389	92	338	443	456	579	887
	US-10-080-960-19	US-09-815-242-5584	US-09-815-242-12456	US-09-813-153-163	US-09-862-027-37	US-10-413-536-4	US-09-738-626-3805	US-09-815-242-5626	US-09-815-242-12247	US-09-815-242-12291	US-09-815-242-12045	US-10-353-929-49	US-10-286-264-106	US-10-007-280A-214	US-09-764-868-895	US-09-925-297-482	US-09-895-913A-126	US-10-156-761-9148	US-09-934-455-326	US-10-021-955-82	US-10-021-955-86	US-10-021-955-87	US-10-021-955-79	US-10-098-916A-11	US-10-106-698-7618	US-09-769-787-146	US-10-205-219-91	US-09-925-300-1585	US-10-156-761-8655	US-09-815-242-13179
	Sequence 19, Appl1	Sequence 5584, Ap	Sequence 12456, A	Sequence 163, Appl	Sequence 37, Appl	Sequence 4, Appl1	Sequence 3805, Ap	Sequence 5626, Ap	Sequence 12247, A	Sequence 12291, A	Sequence 12045, A	Sequence 49, Appl1	Sequence 106, Appl	Sequence 214, App	Sequence 895, App	Sequence 482, App	Sequence 126, App	Sequence 9148, Ap	Sequence 326, App	Sequence 82, Appl	Sequence 86, Appl	Sequence 87, Appl	Sequence 79, Appl	Sequence 11, Appl	Sequence 7618, Ap	Sequence 146, App	Sequence 91, Appl	Sequence 1595, Ap	Sequence 8655, Ap	Sequence 13179, A

ALIGNMENTS

RESULT 1
US-09-740-274-4
Sequence 4, Application US/09740274
Patent No. US20020031826A1
GENERAL INFORMATION:
APPLICANT: Nichols, Scott E.
TITLE OR INVENTION: Glucan-containing Compositions and Paper
FILE REFERENCE: 0357CRD
CURRENT APPLICATION NUMBER: US/09/740,274
CURRENT FILING DATE: 2000-12-19
PRIOR APPLICATION NUMBER: 09/210,361
PRIOR FILING DATE: 1998-12-11
PRIOR APPLICATION NUMBER: 09/007,999
PRIOR FILING DATE: 1998-01-16
PRIOR APPLICATION NUMBER: 08/478,704
PRIOR FILING DATE: 1995-06-07
PRIOR APPLICATION NUMBER: 09/009,620
PRIOR FILING DATE: 1998-01-20
PRIOR APPLICATION NUMBER: 08/485,243
PRIOR FILING DATE: 1995-06-07
PRIOR APPLICATION NUMBER: 09/008,172
PRIOR FILING DATE: 1998-01-16
PRIOR APPLICATION NUMBER: 08/482,711
PRIOR FILING DATE: 1995-06-07
NUMBER OF SEQ ID NOS: 6
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 4
LENGTH: 1375
TYPE: PRT
ORGANISM: streptococcus mutans
US-09-740-274-4
Query Match 100.0%; Score 110; DB 9; Length 1375;
Best local Similarity 100.0%; Pred. No. 3, 6e-09;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 VPSYSTRADSEVQDLIRNII 22
|||||

Db 578 VPSYSFIRAHDSFVODLIRNII 599

RESULT 2

US-09-740-274-2
; Sequence 2, Application US/09740274
; Patent No. US20020031826A1
; GENERAL INFORMATION:
; APPLICANT: Nichols, Scott E.
; TITLE OF INVENTION: Glucan-containing Compositions and Paper
; FILE REFERENCE: 0357CRD
; CURRENT APPLICATION NUMBER: US/09/740,274
; CURRENT FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: 09/210,361
; PRIOR FILING DATE: 1998-12-11
; PRIOR APPLICATION NUMBER: 09/007,999
; PRIOR FILING DATE: 1998-01-16
; PRIOR APPLICATION NUMBER: 08/478,704
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: 09/009,620
; PRIOR FILING DATE: 1998-01-20
; PRIOR APPLICATION NUMBER: 08/485,243
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: 09/008,172
; PRIOR FILING DATE: 1998-01-16
; PRIOR APPLICATION NUMBER: 08/482,711
; PRIOR FILING DATE: 1995-06-07
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 1475
; TYPE: PRT
; ORGANISM: Streptococcus mutans
US-09-740-274-2

Query Match 90.0%; Score 99; DB 9; Length 1475;
Best Local Similarity 90.9%; Pred. No. 2.5e-07;
Matches 20; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 VPSYSFIRAHDSFVODLIRNII 22

Db 552 VPSYSFIRAHDSFVODLIRNII 573

RESULT 3
US-09-995-749A-11
; Sequence 11, Application US/09995749A
; Patent No. US2002015568A1
; GENERAL INFORMATION:
; APPLICANT: VAN GEEL-SCHUTTEN, GERRITDINA HENDRIKA
; APPLICANT: DIKHUIZEN, LOBBERT
; APPLICANT: RAHAOUI, HAKIM
; APPLICANT: LEER, ROBERT-JAN
; TITLE OF INVENTION: NOVEL GLUCOSYLTRANSFERASES
; FILE REFERENCE: BO43388-CIP
; CURRENT APPLICATION NUMBER: US/09/995,749A
; CURRENT FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: 09/604,957
; PRIOR FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: EPO 00201871.1
; PRIOR FILING DATE: 2000-05-25
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 11
; LENGTH: 522
; TYPE: PRT
; ORGANISM: Leuconostoc mesenteroides
US-09-995-749A-11

Query Match 74.5%; Score 82; DB 10; Length 522;
Best Local Similarity 63.6%; Pred. No. 4.7e-05;
Matches 14; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

QY 1 VPSYSFIRAHDSFVODLIRNII 22

Db 146 IPNYSFVRAHDNSQDIOIAL 167

RESULT 4
US-09-995-749A-13
; Sequence 13, Application US/09995749A
; Patent No. US2002015568A1
; GENERAL INFORMATION:
; APPLICANT: VAN GEEL-SCHUTTEN, GERRITDINA HENDRIKA
; APPLICANT: DIKHUIZEN, LOBBERT
; APPLICANT: RAHAOUI, HAKIM
; APPLICANT: LEER, ROBERT-JAN
; TITLE OF INVENTION: NOVEL GLUCOSYLTRANSFERASES
; FILE REFERENCE: BO43388-CIP
; CURRENT APPLICATION NUMBER: US/09/995,749A
; CURRENT FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: 09/604,957
; PRIOR FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: EPO 00201871.1
; PRIOR FILING DATE: 2000-05-25
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 13
; LENGTH: 535
; TYPE: PRT
; ORGANISM: Lactobacillus reuteri
US-09-995-749A-13

Query Match 69.1%; Score 76; DB 10; Length 535;
Best Local Similarity 59.1%; Pred. No. 0.00046;
Matches 13; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY 1 VPSYSFIRAHDSFVODLIRNII 22

Db 144 IPNYSFVRAHDNSQDIOIAL 165

RESULT 5
US-09-995-749A-2
; Sequence 2, Application US/09995749A
; Patent No. US2002015568A1
; GENERAL INFORMATION:
; APPLICANT: VAN GEEL-SCHUTTEN, GERRITDINA HENDRIKA
; APPLICANT: DIKHUIZEN, LOBBERT
; APPLICANT: RAHAOUI, HAKIM
; APPLICANT: LEER, ROBERT-JAN
; TITLE OF INVENTION: NOVEL GLUCOSYLTRANSFERASES
; FILE REFERENCE: BO43388-CIP
; CURRENT APPLICATION NUMBER: US/09/995,749A
; CURRENT FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: 09/604,957
; PRIOR FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: EPO 00201871.1
; PRIOR FILING DATE: 2000-05-25
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 1781
; TYPE: PRT
; ORGANISM: Lactobacillus reuteri
US-09-995-749A-2

Query Match 69.1%; Score 76; DB 10; Length 1781;
Best Local Similarity 59.1%; Pred. No. 0.0018;
Matches 13; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY 1 VPSYSFIRAHDSFVODLIRNII 22

Db 1123 IPNYSFVRAHDNSQDIOIAL 1144

```
RESULT 6
US-09-995-749A-12
; Sequence 12, Application US/09995749A
; Patent No. US2002015568A1
; GENERAL INFORMATION:
; APPLICANT: VAN GEEL-SCHUTTEN, GERRITDINA HENDRIKA
; APPLICANT: DIJKHUIZEN, LOBBERT
; APPLICANT: RAHAOUI, HAKIM
; APPLICANT: LEER, ROBERT-JAN
; TITLE OF INVENTION: NOVEL GLUCOSYLTRANSFERASES
; FILE REFERENCE: BO43388-CIP
; CURRENT APPLICATION NUMBER: US/09/995,749A
; PRIOR FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: 09/604,957
; PRIOR FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: EPO 00201871.1
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 12
; LENGTH: 584
; TYPE: PRT
; ORGANISM: Leuconostoc mesenteroides
US-09-995-749A-12

Query Match          64.5%; Score 71; DB 10; Length 584;
Best Local Similarity 63.2%; Pred. No. 0.0034;
Matches 12; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY      1 VPSYFIRAHSEVQDLIRNI 19
Db      167 IPNYSFVRAHDYDADQPIR 185

RESULT 7
US-09-995-749A-10
; Sequence 10, Application US/09995749A
; Patent No. US2002015568A1
; GENERAL INFORMATION:
; APPLICANT: VAN GEEL-SCHUTTEN, GERRITDINA HENDRIKA
; APPLICANT: DIJKHUIZEN, LOBBERT
; APPLICANT: RAHAOUI, HAKIM
; APPLICANT: LEER, ROBERT-JAN
; TITLE OF INVENTION: NOVEL GLUCOSYLTRANSFERASES
; FILE REFERENCE: BO43388-CIP
; CURRENT APPLICATION NUMBER: US/09/995,749A
; PRIOR FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: 09/604,957
; PRIOR FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: EPO 00201871.1
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
; LENGTH: 545
; TYPE: PRT
; ORGANISM: Streptococcus mutans
US-09-995-749A-10

Query Match          61.8%; Score 68; DB 10; Length 545;
Best Local Similarity 70.0%; Pred. No. 0.0098;
Matches 14; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY      3 SYSFIRAHSEVQDLIRNI 22
Db      156 NYIFIRAHSEVQTVIAKII 175

RESULT 8
US-09-740-274-6
; Sequence 6, Application US/09740274
; Patent No. US20020031826A1
; GENERAL INFORMATION:
; APPLICANT: NICHOLS, Scott E.
; TITLE OF INVENTION: Glucan-containing Compositions and Paper
; FILE REFERENCE: 0357CRD
; CURRENT APPLICATION NUMBER: US/09/740,274
; PRIOR FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: 09/210,361
; PRIOR FILING DATE: 1998-12-11
; PRIOR APPLICATION NUMBER: 09/007,999
; PRIOR FILING DATE: 1998-01-16
; PRIOR APPLICATION NUMBER: 08/478,704
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: 09/009,620
; PRIOR FILING DATE: 1998-01-20
; PRIOR APPLICATION NUMBER: 08/485,243
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: 09/008,172
; PRIOR FILING DATE: 1998-01-16
; PRIOR APPLICATION NUMBER: 08/482,711
; PRIOR FILING DATE: 1995-06-07
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 6
; LENGTH: 1430
; TYPE: PRT
; ORGANISM: Streptococcus mutans
US-09-740-274-6

Query Match          61.8%; Score 68; DB 9; Length 1430;
Best Local Similarity 70.0%; Pred. No. 0.029;
Matches 14; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY      3 SYSFIRAHSEVQDLIRNI 22
Db      576 NYIFIRAHSEVQTVIAKII 595

RESULT 9
US-09-557-796-20
; Sequence 20, Application US/09557796
; Publication No. US20030073140A1
; GENERAL INFORMATION:
; APPLICANT: Hocht, James
; APPLICANT: Dattois, Veronique
; TITLE OF INVENTION: METABOLIC SELECTION METHODS
; FILE REFERENCE: 234/191
; CURRENT APPLICATION NUMBER: US/09/557,796
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/172,952
; PRIOR FILING DATE: 1998-10-14
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 20
; LENGTH: 282
; TYPE: PRT
; ORGANISM: YiaJ-Ec
US-09-557-796-20

Query Match          40.0%; Score 44; DB 11; Length 282;
Best Local Similarity 43.5%; Pred. No. 40;
Matches 10; Conservative 3; Mismatches 8; Indels 2; Gaps 1;

QY      2 PSY--SFRHSEVQDLIRNI 22
Db      169 PDYKSYWESHCHIQPLTRNTI 191

RESULT 10
US-10-101-235A-8
; Sequence 8, Application US/10101235A
; Publication No. US20030182669A1
; GENERAL INFORMATION:
; APPLICANT: Rockman, Howard A.
; APPLICANT: Naga Prasad, Sathyamangla V.
```

```

? APPLICANT: Laporte, Stephane A.
? APPLICANT: Barak, Larry S.
? APPLICANT: Caron, Marc G.
? TITLE OR INVENTION: Phosphoinositide 3-Kinase Mediated Inhibition of GPCRs
? FILE REFERENCE: 033072-064
? CURRENT APPLICATION NUMBER: US/10/101,235A
? CURRENT FILING DATE: 2002-03-19
? NUMBER OF SEQ ID NOS: 9
? SOFTWARE: FastSeq for Windows Version 4.0
? SEQ ID NO 8
? LENGTH: 1068
? TYPE: PRT
? ORGANISM: Homo sapiens
? OS-10-101-235A-8

```

Query Match	40.0%	Score 44:	DB 12:	Length 1068;
Best Local Similarity	45.5%	Pred. No.	1.8e+02;	
Matches 10;	Conservative 2;	Mismatches 10;	Indels 0;	Gaps 0;

```

QY      1 VPSYSFIRAHDSVEVDLIRNII 22
          :| | | | | | | | | |
Db      123 MPVCEFDNVKDPEVQDERRNIL 144

```

RESULT 11
US-10-012-542-277
; Sequence 277, Application US/10012542
; Publication No. US20030044851A1

1 APPLICANT: Ruben et al.
 2 TITLE OF INVENTION: 94 Human Secreted Proteins
 3 FILE REFERENCE: P202991
 4 CURRENT APPLICATION NUMBER: US/10/012,542
 5 CURRENT FILING DATE: 2001-12-12
 6 PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/461,355
 7 PRIOR FILING DATE: EARLIER FILING DATE: 1999-12-14
 8 PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/089,507
 9 PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-16
 10 PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/089,508
 11 PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-16
 12 PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/089,509
 13 PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-16
 14 PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/089,510
 15 PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-16
 16 PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/090,112
 17 PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-22
 18 PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/090,113
 19 PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-22

```

1 / NUMBER OF SEQ ID NOS: 532
2 / SOFTWARE: PatentIn Ver. 2.0
3 / SEQ ID NO 277
4 / LENGTH: 347
5 / TYPE: PRT
6 / ORGANISM: Homo sapiens
7 / FEATURE:
8 / NAME/KEY: SITE
9 / LOCATION: (16)
10 / OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
11 / NAME/KEY: SITE
12 / LOCATION: (340)
13 / OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
14 / NAME/KEY: SITE
15 / LOCATION: (341)
16 / OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
17 /
18 /
19 /
20 /
21 /
22 /
23 /
24 /
25 /
26 /
27 /
28 /
29 /
30 /
31 /
32 /
33 /
34 /
35 /
36 /
37 /
38 /
39 /
40 /
41 /
42 /
43 /
44 /
45 /
46 /
47 /
48 /
49 /
50 /
51 /
52 /
53 /
54 /
55 /
56 /
57 /
58 /
59 /
60 /
61 /
62 /
63 /
64 /
65 /
66 /
67 /
68 /
69 /
70 /
71 /
72 /
73 /
74 /
75 /
76 /
77 /
78 /
79 /
80 /
81 /
82 /
83 /
84 /
85 /
86 /
87 /
88 /
89 /
90 /
91 /
92 /
93 /
94 /
95 /
96 /
97 /
98 /
99 /
100 /
101 /
102 /
103 /
104 /
105 /
106 /
107 /
108 /
109 /
110 /
111 /
112 /
113 /
114 /
115 /
116 /
117 /
118 /
119 /
120 /
121 /
122 /
123 /
124 /
125 /
126 /
127 /
128 /
129 /
130 /
131 /
132 /
133 /
134 /
135 /
136 /
137 /
138 /
139 /
140 /
141 /
142 /
143 /
144 /
145 /
146 /
147 /
148 /
149 /
150 /
151 /
152 /
153 /
154 /
155 /
156 /
157 /
158 /
159 /
160 /
161 /
162 /
163 /
164 /
165 /
166 /
167 /
168 /
169 /
170 /
171 /
172 /
173 /
174 /
175 /
176 /
177 /
178 /
179 /
180 /
181 /
182 /
183 /
184 /
185 /
186 /
187 /
188 /
189 /
190 /
191 /
192 /
193 /
194 /
195 /
196 /
197 /
198 /
199 /
200 /
201 /
202 /
203 /
204 /
205 /
206 /
207 /
208 /
209 /
210 /
211 /
212 /
213 /
214 /
215 /
216 /
217 /
218 /
219 /
220 /
221 /
222 /
223 /
224 /
225 /
226 /
227 /
228 /
229 /
230 /
231 /
232 /
233 /
234 /
235 /
236 /
237 /
238 /
239 /
240 /
241 /
242 /
243 /
244 /
245 /
246 /
247 /
248 /
249 /
250 /
251 /
252 /
253 /
254 /
255 /
256 /
257 /
258 /
259 /
260 /
261 /
262 /
263 /
264 /
265 /
266 /
267 /
268 /
269 /
270 /
271 /
272 /
273 /
274 /
275 /
276 /
277 /
278 /
279 /
280 /
281 /
282 /
283 /
284 /
285 /
286 /
287 /
288 /
289 /
290 /
291 /
292 /
293 /
294 /
295 /
296 /
297 /
298 /
299 /
300 /
301 /
302 /
303 /
304 /
305 /
306 /
307 /
308 /
309 /
310 /
311 /
312 /
313 /
314 /
315 /
316 /
317 /
318 /
319 /
320 /
321 /
322 /
323 /
324 /
325 /
326 /
327 /
328 /
329 /
330 /
331 /
332 /
333 /
334 /
335 /
336 /
337 /
338 /
339 /
340 /
341 /
342 /
343 /
344 /
345 /
346 /
347 /
348 /
349 /
350 /
351 /
352 /
353 /
354 /
355 /
356 /
357 /
358 /
359 /
360 /
361 /
362 /
363 /
364 /
365 /
366 /
367 /
368 /
369 /
370 /
371 /
372 /
373 /
374 /
375 /
376 /
377 /
378 /
379 /
380 /
381 /
382 /
383 /
384 /
385 /
386 /
387 /
388 /
389 /
390 /
391 /
392 /
393 /
394 /
395 /
396 /
397 /
398 /
399 /
400 /
401 /
402 /
403 /
404 /
405 /
406 /
407 /
408 /
409 /
410 /
411 /
412 /
413 /
414 /
415 /
416 /
417 /
418 /
419 /
420 /
421 /
422 /
423 /
424 /
425 /
426 /
427 /
428 /
429 /
430 /
431 /
432 /
433 /
434 /
435 /
436 /
437 /
438 /
439 /
440 /
441 /
442 /
443 /
444 /
445 /
446 /
447 /
448 /
449 /
450 /
451 /
452 /
453 /
454 /
455 /
456 /
457 /
458 /
459 /
460 /
461 /
462 /
463 /
464 /
465 /
466 /
467 /
468 /
469 /
470 /
471 /
472 /
473 /
474 /
475 /
476 /
477 /
478 /
479 /
480 /
481 /
482 /
483 /
484 /
485 /
486 /
487 /
488 /
489 /
490 /
491 /
492 /
493 /
494 /
495 /
496 /
497 /
498 /
499 /
500 /
501 /
502 /
503 /
504 /
505 /
506 /
507 /
508 /
509 /
510 /
511 /
512 /
513 /
514 /
515 /
516 /
517 /
518 /
519 /
520 /
521 /
522 /
523 /
524 /
525 /
526 /
527 /
528 /
529 /
530 /
531 /
532 /
533 /
534 /
535 /
536 /
537 /
538 /
539 /
540 /
541 /
542 /
543 /
544 /
545 /
546 /
547 /
548 /
549 /
550 /
551 /
552 /
553 /
554 /
555 /
556 /
557 /
558 /
559 /
560 /
561 /
562 /
563 /
564 /
565 /
566 /
567 /
568 /
569 /
570 /
571 /
572 /
573 /
574 /
575 /
576 /
577 /
578 /
579 /
580 /
581 /
582 /
583 /
584 /
585 /
586 /
587 /
588 /
589 /
590 /
591 /
592 /
593 /
594 /
595 /
596 /
597 /
598 /
599 /
600 /
601 /
602 /
603 /
604 /
605 /
606 /
607 /
608 /
609 /
610 /
611 /
612 /
613 /
614 /
615 /
616 /
617 /
618 /
619 /
620 /
621 /
622 /
623 /
624 /
625 /
626 /
627 /
628 /
629 /
630 /
631 /
632 /
633 /
634 /
635 /
636 /
637 /
638 /
639 /
640 /
641 /
642 /
643 /
644 /
645 /
646 /
647 /
648 /
649 /
650 /
651 /
652 /
653 /
654 /
655 /
656 /
657 /
658 /
659 /
660 /
661 /
662 /
663 /
664 /
665 /
666 /
667 /
668 /
669 /
670 /
671 /
672 /
673 /
674 /
675 /
676 /
677 /
678 /
679 /
680 /
681 /
682 /
683 /
684 /
685 /
686 /
687 /
688 /
689 /
690 /
691 /
692 /
693 /
694 /
695 /
696 /
697 /
698 /
699 /
700 /
701 /
702 /
703 /
704 /
705 /
706 /
707 /
708 /
709 /
710 /
711 /
712 /
713 /
714 /
715 /
716 /
717 /
718 /
719 /
720 /
721 /
722 /
723 /
724 /
725 /
726 /
727 /
728 /
729 /
730 /
731 /
732 /
733 /
734 /
735 /
736 /
737 /
738 /
739 /
740 /
741 /
742 /
743 /
744 /
745 /
746 /
747 /
748 /
749 /
750 /
751 /
752 /
753 /
754 /
755 /
756 /
757 /
758 /
759 /
760 /
761 /
762 /
763 /
764 /
765 /
766 /
767 /
768 /
769 /
770 /
771 /
772 /
773 /
774 /
775 /
776 /
777 /
778 /
779 /
780 /
781 /
782 /
783 /
784 /
785 /
786 /
787 /
788 /
789 /
790 /
791 /
792 /
793 /
794 /
795 /
796 /
797 /
798 /
799 /
800 /
801 /
802 /
803 /
804 /
805 /
806 /
807 /
808 /
809 /
810 /
811 /
812 /
813 /
814 /
815 /
8
```

Query Match Similarity 39.1%; Score 43; DB 15; Length 347;
Best Local Similarity 35.3%; Pred. No. 74;
Matches 6; Conservative 7; Mismatches 4; Indels 0; Gaps 0;

```
QY      4 YSFRAHDSEVQDLIRN 20
        : : : : : : : :
Db     103 HKYMRSHKAESSELVRN 119
```

RESULT 12
US-10-156-761-11791
; Sequence 11791, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:

1 APPLICANT: OMIURA, SATOSHI
 2 APPLICANT: IKEDA, HARUO
 3 APPLICANT: ISHIKAWA, JUN
 4 APPLICANT: HORIKAWA, HIROSHI
 5 APPLICANT: SHIBA, TADAYOSHI
 6 APPLICANT: SARAKI, YOSHIYUKI
 7 APPLICANT: HATTORI, MASAHIRA
 8 TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
 9 FILE REFERENCE: 249-262
 10
 11 CURRENT APPLICATION NUMBER: US/10/156,761
 12 CURRENT FILING DATE: 2002-05-29
 13 PRIOR APPLICATION NUMBER: JP 2001-204089
 14 PRIOR FILING DATE: 2001-05-30
 15 PRIOR APPLICATION NUMBER: JP 2001-272697
 16 PRIOR FILING DATE: 2001-08-02
 17 NUMBER OF SEQ. ID NOS: 15109
 18 SEQ. ID NO 11791

; ORGANISM: Streptomyces avermitilis
US-10-156-761-11791

Query Match	39.1%	Score 43	DB 15	Length 568
Best Local Similarity	47.4%	Pred. NC	1.3e+02	
Matches	9	Conservative	3	Mismatches 7
				Indels 0
				Gaps 0

```
QY      2 PSYSFIRAHDSVQDLIRN 20
      ||| | : | : | ||
Db      456 PSYRFLGPVDTAVGELTRN 474
```

RESULT 13
US-10-177-293-470
; Sequence 470, Application US/10177293
; Publication No. US20030124128A1

GENERAL INFORMATION:
APPLICANT: Lillie, James
APPLICANT: Ghatt, Karen
APPLICANT: Zhao, Xumei
APPLICANT: Gannavarapu, Manjula
APPLICANT: Kamatkar, Shubhangt
APPLICANT: Mertens, Maureen
APPLICANT: Myer, Vlc
APPLICANT: Wang, Youzhen
APPLICANT: Xu, Yongyao
APPLICANT: Hoersch, Sebastian
APPLICANT: Monahan, John
APPLICANT: Meyers, Rachel E.
APPLICANT: Bast Jr., Robert C.
APPLICANT: Hortobaagi, Gabriel N.
APPLICANT: Pusztai, Iajos
APPLICANT: Meric, Funda
APPLICANT: Sahin, Aysegul
APPLICANT: Mills, Gordon B.
TITLE OF INVENTION: COMPOSITIONS, KITS, AND METHODS FOR IDENTIFICATION, ASSESSMENT
TITLE OF INVENTION: PREVENTION, AND THERAPY OF BREAST CANCER
FILE REFERENCE: MRI-038
CURRENT APPLICATION NUMBER: US/10/177,293
CURRENT FILING DATE: 2002-06-21
PRIOR APPLICATION NUMBER: US 60/299,887
PRIOR FILING DATE: 2001-06-21
PRIOR APPLICATION NUMBER: US 60/301,572
PRIOR FILING DATE: 2001-06-27
PRIOR APPLICATION NUMBER: US 60/306,501
PRIOR FILING DATE: 2001-07-18
PRIOR APPLICATION NUMBER: US 60/325,002

```

; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US 60/362,585
; PRIOR FILING DATE: 2002-03-05
; PRIOR APPLICATION NUMBER: US 60/xxx,xxx
; PRIOR FILING DATE: 2002-05-14
; NUMBER OF SEQ ID NOS: 506
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 470
; LENGTH: 2541
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-177-293-470
```

```

Query Match          39.1%; Score 43; DB 15; Length 2541;
Best Local Similarity 38.1%; Pred. No. 7.2e+02;
Matches 8; Conservative 4; Mismatches 9; Indels 0; Gaps 0;
```

```
QY      2 PSYSFIRAHSEVODLIRNII 22
Db      1628 PSWSVLAGHSRTVSISIKKLI 1648
```

```

RESULT 14
US-09-815-242-5358
; Sequence 5358, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; TITLE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5358
; LENGTH: 1213
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-815-242-5358
```

```

Query Match          37.7%; Score 41.5; DB 9; Length 1213;
Best Local Similarity 43.5%; Pred. No. 5.4e+02;
Matches 10; Conservative 6; Mismatches 2; Indels 5; Gaps 1;
```

```
QY      5 SFIKRAHSEVOD-----LIRNII 22
Db      615 SFLKAIKPLQDIYLVGLMRSVI 637
```

```

RESULT 15
US-09-815-242-12523
```

```

; Sequence 12523, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; TITLE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12523
; LENGTH: 1217
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-815-242-12523
```

```

Query Match          37.7%; Score 41.5; DB 9; Length 1217;
Best Local Similarity 43.5%; Pred. No. 5.5e+02;
Matches 10; Conservative 6; Mismatches 2; Indels 5; Gaps 1;
```

```
QY      5 SFIKRAHSEVOD-----LIRNII 22
Db      615 SFLKAIKPLQDIYLVGLMRSVI 637
```

```

Search completed: November 13, 2003, 10:29:04
Job time : 28.7773 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 13, 2003, 09:31:40 ; Search time 14.2844 Seconds
(without alignments)
65.165 Million cell updates/sec

Title: US-09-290-049A-16
Perfect score: 107
Sequence: 1 MANYIFIRAHDSVQVIANKII 22

Scoring table:
BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 4210858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

- Database :
- 1: /cgn2_6/ptodata/2/iaa/5A.COMB.dep.*
 - 2: /cgn2_6/ptodata/2/iaa/5B.COMB.dep.*
 - 3: /cgn2_6/ptodata/2/iaa/6A.COMB.dep.*
 - 4: /cgn2_6/ptodata/2/iaa/6B.COMB.dep.*
 - 5: /cgn2_6/ptodata/2/iaa/PCTUS.COMB.dep.*
 - 6: /cgn2_6/ptodata/2/iaa/backfile1.dep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	107	100.0	545	4	US-09-604-957-4
2	107	100.0	1430	3	US-09-008-172-2
3	107	100.0	1430	3	US-09-210-361-6
4	107	100.0	1430	4	US-09-740-274-6
5	89	83.2	1577	2	US-08-793-824-2
6	86	80.4	521	4	US-09-604-957-5
7	72	67.3	1475	3	US-09-007-999-2
8	72	67.3	1475	3	US-09-210-361-2
9	72	67.3	1475	4	US-09-740-274-2
10	68	63.6	1375	3	US-09-210-361-4
11	68	63.6	1375	4	US-09-740-274-4
12	55	51.4	584	4	US-09-604-957-6
13	53	49.5	2057	4	US-09-499-203-2
14	50	46.7	535	4	US-09-604-957-7
15	50	46.7	1278	4	US-09-604-957-3
16	44	41.1	183	4	US-09-107-532A-5511
17	40.5	37.9	279	4	US-09-107-532A-6807
18	40	37.4	871	4	US-09-585-858-57
19	40	37.4	871	4	US-09-585-858-36
20	40	37.4	2270	4	US-09-581-909-3
21	39	36.4	182	4	US-09-338-907-133
22	39	36.4	182	4	US-09-218-207-133
23	39	36.4	185	3	US-09-338-907-136
24	39	36.4	185	4	US-09-218-207-136
25	39	36.4	228	3	US-09-338-907-70
26	39	36.4	228	4	US-09-218-207-70
27	39	36.4	300	3	US-09-338-907-135

28	39	36.4	300	4	US-09-218-207-135	Sequence 135, App
29	39	36.4	315	3	US-09-338-907-134	Sequence 134, App
30	39	36.4	315	4	US-09-218-207-134	Sequence 134, App
31	39	36.4	353	2	US-08-996-306-4	Sequence 4, Appli
32	39	36.4	353	3	US-09-338-907-4	Sequence 4, Appli
33	39	36.4	353	4	US-09-218-207-4	Sequence 4, Appli
34	39	36.4	364	2	US-08-996-306-5	Sequence 5, Appli
35	39	36.4	364	3	US-09-338-907-5	Sequence 5, Appli
36	39	36.4	364	4	US-09-218-207-5	Sequence 5, Appli
37	38	35.5	121	4	US-09-134-001C-3077	Sequence 3077, Ap
38	38	35.5	146	2	US-08-647-960-8	Sequence 8, Appli
39	38	35.5	257	3	US-09-287-097-2	Sequence 2, Appli
40	38	35.5	260	2	US-08-685-992-4	Sequence 4, Appli
41	38	35.5	260	2	US-08-685-992-25	Sequence 25, Appli
42	38	35.5	260	2	US-09-144-925-4	Sequence 4, Appli
43	38	35.5	260	2	US-09-144-925-25	Sequence 25, Appli
44	38	35.5	261	2	US-08-685-992-3	Sequence 3, Appli
45	38	35.5	261	2	US-09-144-925-3	Sequence 3, Appli

ALIGNMENTS

```
RESULT 1
US-09-604-957-4
; Sequence 4, Application US/09604957
; Patent No. 6486314
; GENERAL INFORMATION:
; APPLICANT: VAN GEEL-SCHUTTEN, GERITDINA HENDRIKA
; APPLICANT: DIJKHITZEN, LUBBERT
; APPLICANT: RAHMOU, HAKIM
; APPLICANT: LEBER, ROBERT-JAN
; TITLE OF INVENTION: NOVEL GLUCOSYLTRANSFERASE AND NOVEL GLUCAN
; FILE REFERENCE: BO 43388
; CURRENT APPLICATION NUMBER: US/09/604,957
; CURRENT FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: 00201871.1
; PRIOR FILING DATE: 2000-05-25
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 545
; TYPE: PRT
; ORGANISM: Streptococcus mutans
; US-09-604-957-4

Query Match      100.0%; Score 107; DB 4; Length 545;
Best Local Similarity 100.0%; Pred. No. 4, 4e-11;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy      1 MANYIFIRAHDSVQVIANKII 22
Db      154 MANYIFIRAHDSVQVIANKII 175

RESULT 2
US-09-008-172-2
; Sequence 2, Application US/09008172
; Patent No. 6127602
; GENERAL INFORMATION:
; APPLICANT: Nichols, Scott E.
; TITLE OF INVENTION: Substitutes for Modified Starch and
; FILE REFERENCE: Lateges in Paper Manufacture
; CURRENT APPLICATION NUMBER: US/09/008,172
; CURRENT FILING DATE: 1998-01-16
; EARLIER APPLICATION NUMBER: 08/482,711
; EARLIER FILING DATE: 1995-06-07
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 1430
; TYPE: PRT
```

```

; ORGANISM: Streptococcus mutans
US-09-008-172-2
Query Match          100.0%; Score 107; DB 3; Length 1430;
Best Local Similarity 100.0%; Pred. No. 1.4e-10;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 MANYIFIRAHDSVQTVIAKII 22
      |||:|||||:|||||:|||||:
Db      574 MANYIFIRAHDSVQTVIAKII 595

RESULT 3
US-09-210-361-6
; Sequence 6, Application US/09210361
; Patent No. 6284479
; GENERAL INFORMATION:
; APPLICANT: Nichols, Scott E.
; TITLE OF INVENTION: Substitutes for Modified Starches and
; FILE REFERENCE: 0357CR
; CURRENT APPLICATION NUMBER: US/09/210,361
; EARLIER FILING DATE: 1998-12-11
; EARLIER APPLICATION NUMBER: 09/007,999
; EARLIER FILING DATE: 1998-01-16
; EARLIER APPLICATION NUMBER: 08/478,704
; EARLIER FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 09/009,620
; EARLIER FILING DATE: 1998-01-20
; EARLIER APPLICATION NUMBER: 08/485,243
; EARLIER FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 09/008,172
; EARLIER FILING DATE: 1998-01-16
; EARLIER APPLICATION NUMBER: 08/482,711
; EARLIER FILING DATE: 1995-06-07
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 6
; LENGTH: 1430
; TYPE: PRF
; ORGANISM: streptococcus mutans
US-09-210-361-6

Query Match          100.0%; Score 107; DB 3; Length 1430;
Best Local Similarity 100.0%; Pred. No. 1.4e-10;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 MANYIFIRAHDSVQTVIAKII 22
      |||:|||||:|||||:|||||:
Db      574 MANYIFIRAHDSVQTVIAKII 595

RESULT 4
US-09-740-274-6
; Sequence 6, Application US/09740274
; Patent No. 6465203
; GENERAL INFORMATION:
; APPLICANT: Nichols, Scott E.
; TITLE OF INVENTION: Glucan-containing Compositions and Paper
; FILE REFERENCE: 0357CRD
; CURRENT APPLICATION NUMBER: US/09/740,274
; EARLIER FILING DATE: 2000-12-19
; EARLIER APPLICATION NUMBER: 09/210,361
; EARLIER FILING DATE: 1998-12-11
; EARLIER APPLICATION NUMBER: 09/007,999
; EARLIER FILING DATE: 1998-01-16
; EARLIER APPLICATION NUMBER: 08/478,704
; EARLIER FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 09/009,620
; EARLIER FILING DATE: 1998-01-20
; EARLIER APPLICATION NUMBER: 08/485,243
; EARLIER FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 09/008,172
```

```

; PRIOR FILING DATE: 1998-01-16
; PRIOR APPLICATION NUMBER: 08/482,711
; PRIOR FILING DATE: 1995-06-07
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 6
; LENGTH: 1430
; TYPE: PRF
; ORGANISM: streptococcus mutans
US-09-740-274-6

Query Match          100.0%; Score 107; DB 4; Length 1430;
Best Local Similarity 100.0%; Pred. No. 1.4e-10;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 MANYIFIRAHDSVQTVIAKII 22
      |||:|||||:|||||:|||||:
Db      574 MANYIFIRAHDSVQTVIAKII 595

RESULT 5
US-08-793-824-2
; Sequence 2, Application US/08793824
; Patent No. 5981838
; GENERAL INFORMATION:
; APPLICANT: Simpson, Christine Lynn
; APPLICANT: Giffard, Philip Morrison
; APPLICANT: Jacques, Nicholas Anthony
; TITLE OF INVENTION: Genetic Manipulation of Plants to
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESS: Griffith Hack & Co
; STREET: Level 8, 168 Walker Street
; CITY: No. 5981838th Sydney
; STATE: New South Wales
; COUNTRY: Australia
; ZIP: 2060
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/793,824
; FILING DATE:
; CLASSIFICATION: 900
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: AU PM7643
; FILING DATE: 24-AUG-1994
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 61 2 9957 5944
; TELEFAX: 61 2 957 6288
; TELEX: 26547
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1577 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: not relevant
; MOLECULE TYPE: protein
; ORIGINAL SOURCE:
; ORGANISM: Streptococcus salivarius
US-08-793-824-2

Query Match          83.2%; Score 89; DB 2; Length 1577;
Best Local Similarity 81.0%; Pred. No. 2.7e-07;
Matches 17; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY      2 ANYIFIRAHDSVQTVIAKII 22
      |||:|||||:|||||:|||||:
Db      660 ANYIFIRAHDSVQTVIAKII 680
```

RESULT 6
US-09-604-957-5
; Sequence 5, Application US/09604957
; Patent No. 6486314
; GENERAL INFORMATION:
; APPLICANT: VAN GEEL-SCHUTTEN, GERRITJINA HENDRIKA
; APPLICANT: DIKHUIZEN, LUBBERT
; APPLICANT: RAHAOUI, HAKIM
; APPLICANT: LEER, ROBERT-JAN
; TITLE OF INVENTION: NOVEL GLUCOSYLTRANSFERASE AND NOVEL GLUCAN
; FILE REFERENCE: BO 4338
; CURRENT APPLICATION NUMBER: US/09/604,957
; CURRENT FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: 00201871.1
; PRIOR FILING DATE: 2000-05-25
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 523
; TYPE: PRT
; ORGANISM: Leuconostoc mesenteroides
US-09-604-957-5

Query Match 80.4%; Score 86; DB 4; Length 523;
Best Local Similarity 80.0%; Pred. No. 2.5e-07;
Matches 16; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 3 NYIFIRADSEVQTVIAKII 22
Db 148 NYSFIRADSEVQTVIAQIV 167

RESULT 7
US-09-007-999-2
; Sequence 2, Application US/09007999
; Patent No. 6087559
; GENERAL INFORMATION:
; APPLICANT: Nichols, Scott E.
; TITLE OF INVENTION: Substitutes for Modified Starch and
; FILE REFERENCE: 0356D
; CURRENT APPLICATION NUMBER: US/09/007,999
; CURRENT FILING DATE: 1998-01-16
; EARLIER APPLICATION NUMBER: 08/478,704
; EARLIER FILING DATE: 1995-06-07
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 1475
; TYPE: PRT
; ORGANISM: Streptococcus mutans
US-09-007-999-2

Query Match 67.3%; Score 72; DB 3; Length 1475;
Best Local Similarity 75.0%; Pred. No. 0.00029;
Matches 15; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 3 NYIFIRADSEVQTVIAKII 22
Db 554 SYSFIRADSEVQDLIADI 573

RESULT 8
US-09-210-361-2
; Sequence 2, Application US/09210361
; Patent No. 6284479
; GENERAL INFORMATION:
; APPLICANT: Nichols, Scott E.
; TITLE OF INVENTION: Substitutes for Modified Starches and
; FILE REFERENCE: 0357CR
; CURRENT APPLICATION NUMBER: US/09/210,361

; CURRENT FILING DATE: 1998-12-11
; EARLIER APPLICATION NUMBER: 09/007,999
; EARLIER FILING DATE: 1998-01-16
; EARLIER APPLICATION NUMBER: 08/478,704
; EARLIER FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 09/009,620
; EARLIER FILING DATE: 1998-01-20
; EARLIER APPLICATION NUMBER: 08/485,243
; EARLIER FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 09/008,172
; EARLIER FILING DATE: 1998-01-16
; EARLIER APPLICATION NUMBER: 08/482,711
; EARLIER FILING DATE: 1995-06-07
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 1475
; TYPE: PRT
; ORGANISM: Streptococcus mutans
US-09-210-361-2

Query Match 67.3%; Score 72; DB 3; Length 1475;
Best Local Similarity 75.0%; Pred. No. 0.00029;
Matches 15; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 3 NYIFIRADSEVQTVIAKII 22
Db 554 SYSFIRADSEVQDLIADI 573

RESULT 9
US-09-740-274-2
; Sequence 2, Application US/09740274
; Patent No. 6465203
; GENERAL INFORMATION:
; APPLICANT: Nichols, Scott E.
; TITLE OF INVENTION: Glucan-containing Compositions and Paper
; FILE REFERENCE: 0357CRD
; CURRENT APPLICATION NUMBER: US/09/740,274
; CURRENT FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: 09/210,361
; PRIOR FILING DATE: 1998-12-11
; PRIOR APPLICATION NUMBER: 09/007,999
; PRIOR FILING DATE: 1998-01-16
; PRIOR APPLICATION NUMBER: 08/478,704
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: 09/009,620
; PRIOR FILING DATE: 1998-01-20
; PRIOR APPLICATION NUMBER: 08/485,243
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: 09/008,172
; PRIOR FILING DATE: 1998-01-16
; PRIOR APPLICATION NUMBER: 08/482,711
; PRIOR FILING DATE: 1995-06-07
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 1475
; TYPE: PRT
; ORGANISM: Streptococcus mutans
US-09-740-274-2

Query Match 67.3%; Score 72; DB 4; Length 1475;
Best Local Similarity 75.0%; Pred. No. 0.00029;
Matches 15; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 3 NYIFIRADSEVQTVIAKII 22
Db 554 SYSFIRADSEVQDLIADI 573

RESULT 10
US-09-210-361-4

Query Match	63.6%;	Score 68;	DB 4;	Length 1375;
Best Local Similarity	70.0%;	Pred. No. 0.0014;		
Matches 14;	Conservative 2;	Mismatches 4;	Indels 0;	Gaps 0

RESULT 14
US-09-604-957-7
; Sequence 7, Application US/09604957
; Patent No. 6486314
; GENERAL INFORMATION:

```

; APPLICANT: VAN GEEL-SCHUTTEN, GERRITDINA HENDRIKA
; APPLICANT: DIJKHUIZEN, LOBBERT
; APPLICANT: RAHAOUI, HAKIM
; APPLICANT: LEER, ROBERT-JAN
; TITLE OF INVENTION: NOVEL GLUCOSYLTRANSFERASE AND NOVEL GLUCAN
; FILE REFERENCE: BO 43388
; CURRENT APPLICATION NUMBER: US/09/604,957
; CURRENT FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: 00201871.1
; PRIOR FILING DATE: 2000-05-25
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 7
; LENGTH: 535
; TYPE: PRT
; ORGANISM: Lactobacillus reuteri
US-09-604-957-7

```

```

Query Match          46.7%; Score 50; DB 4; Length 535;
Best Local Similarity 50.0%; Pred. No. 0.81;
Matches 10; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

```

```

QY      3 NYIFRAHDSVQTVIAKII 22
Db      146 NYSFVRAHDNNSQDQIONAI 165

```

```

RESULT 15
US-09-604-957-3
; Sequence 3, Application US/09604957
; Patent No. 6486314
; GENERAL INFORMATION:
; APPLICANT: VAN GEEL-SCHUTTEN, GERRITDINA HENDRIKA
; APPLICANT: DIJKHUIZEN, LOBBERT
; APPLICANT: RAHAOUI, HAKIM
; APPLICANT: LEER, ROBERT-JAN
; TITLE OF INVENTION: NOVEL GLUCOSYLTRANSFERASE AND NOVEL GLUCAN
; FILE REFERENCE: BO 43388
; CURRENT APPLICATION NUMBER: US/09/604,957
; CURRENT FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: 00201871.1
; PRIOR FILING DATE: 2000-05-25
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 3
; LENGTH: 1278
; TYPE: PRT
; ORGANISM: Lactobacillus reuteri.
US-09-604-957-3

```

```

Query Match          46.7%; Score 50; DB 4; Length 1278;
Best Local Similarity 50.0%; Pred. No. 2.3;
Matches 10; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

```

```

QY      3 NYIFRAHDSVQTVIAKII 22
Db      622 NYSFVRAHDNNSQDQIONAI 641

```

```

Search completed: November 13, 2003, 09:48:02
Job time : 15.2844 secs

```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 13, 2003, 09:45:40 ; Search time 28.7773 Seconds
(without alignments)
139.565 Million cell updates/sec

Title: US-09-290-049A-16

Perfect score: 107

Sequence: 1 MANYIFIRAHDSHVQTVIAKII 22

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 666188 seqs, 182559486 residues

Total number of hits satisfying chosen parameters: 666188

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

1: /cgn2_6/ptodata/2/pubppaa/US07_PUBCOMB.pep:*

2: /cgn2_6/ptodata/2/pubppaa/PCT_NEW_PUB.pep:*

3: /cgn2_6/ptodata/2/pubppaa/US06_NEW_PUB.pep:*

4: /cgn2_6/ptodata/2/pubppaa/US06_PUBCOMB.pep:*

5: /cgn2_6/ptodata/2/pubppaa/US07_NEW_PUB.pep:*

6: /cgn2_6/ptodata/2/pubppaa/PCTUS_PUBCOMB.pep:*

7: /cgn2_6/ptodata/2/pubppaa/US08_NEW_PUB.pep:*

8: /cgn2_6/ptodata/2/pubppaa/US08_PUBCOMB.pep:*

9: /cgn2_6/ptodata/2/pubppaa/US09_PUBCOMB.pep:*

10: /cgn2_6/ptodata/2/pubppaa/US09C_PUBCOMB.pep:*

11: /cgn2_6/ptodata/2/pubppaa/US09C_NEW_PUB.pep:*

12: /cgn2_6/ptodata/2/pubppaa/US09_PUBCOMB.pep:*

13: /cgn2_6/ptodata/2/pubppaa/US10_PUBCOMB.pep:*

14: /cgn2_6/ptodata/2/pubppaa/US10C_PUBCOMB.pep:*

15: /cgn2_6/ptodata/2/pubppaa/US10C_NEW_PUB.pep:*

16: /cgn2_6/ptodata/2/pubppaa/US60_NEW_PUB.pep:*

17: /cgn2_6/ptodata/2/pubppaa/US60_PUBCOMB.pep:*

18: /cgn2_6/ptodata/2/pubppaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	107	100.0	545	10	US-09-995-749A-10
2	107	100.0	1430	9	US-09-740-274-6
3	86	80.4	522	9	US-09-995-749A-11
4	72	67.3	1475	9	US-09-740-274-2
5	68	63.6	1375	9	US-09-740-274-4
6	53	49.5	584	10	US-09-995-749A-12
7	50	46.7	535	10	US-09-995-749A-13
8	46	46.7	1781	10	US-09-995-749A-2
9	44	41.1	583	10	US-09-758-269-2
10	42	39.3	327	12	US-10-032-585-7827
11	40	37.4	67	15	US-10-007-280A-214
12	40	37.4	349	15	US-10-270-875-57
13	40	37.4	349	15	US-10-270-878-57
14	40	37.4	349	15	US-10-270-786-57
15	40	37.4	349	15	US-10-270-710-57

16	40	37.4	349	15	US-10-270-859-57	Sequence 57, Appl
17	40	37.4	349	16	US-10-270-846-57	Sequence 56, Appl
18	40	37.4	871	15	US-10-270-875-36	Sequence 36, Appl
19	40	37.4	871	15	US-10-270-878-36	Sequence 36, Appl
20	40	37.4	871	15	US-10-270-786-36	Sequence 36, Appl
21	40	37.4	871	15	US-10-270-710-36	Sequence 36, Appl
22	40	37.4	871	15	US-10-270-859-36	Sequence 36, Appl
23	40	37.4	871	16	US-10-270-846-36	Sequence 36, Appl
24	39	36.4	78	9	US-09-864-761-41675	Sequence 41675, A
25	39	36.4	115	11	US-09-764-891-5098	Sequence 5098, Ap
26	39	36.4	182	10	US-09-901-484A-133	Sequence 133, App
27	39	36.4	182	10	US-09-853-526-133	Sequence 133, App
28	39	36.4	185	10	US-09-901-484A-136	Sequence 136, App
29	39	36.4	185	10	US-09-853-526-136	Sequence 136, App
30	39	36.4	228	10	US-09-901-484A-70	Sequence 70, Appl
31	39	36.4	228	10	US-09-853-526-70	Sequence 70, Appl
32	39	36.4	273	9	US-09-815-242-11100	Sequence 11100, A
33	39	36.4	300	10	US-09-901-484A-135	Sequence 135, App
34	39	36.4	300	10	US-09-853-526-135	Sequence 135, App
35	39	36.4	315	10	US-09-901-484A-134	Sequence 134, App
36	39	36.4	315	10	US-09-853-526-134	Sequence 134, App
37	39	36.4	353	10	US-09-901-484A-4	Sequence 4, Appl
38	39	36.4	353	10	US-09-915-582-59	Sequence 59, Appl
39	39	36.4	353	10	US-09-915-582-79	Sequence 79, Appl
40	39	36.4	353	10	US-09-853-526-4	Sequence 4, Appl
41	39	36.4	353	12	US-10-277-802-59	Sequence 59, Appl
42	39	36.4	353	12	US-10-277-802-79	Sequence 79, Appl
43	39	36.4	364	9	US-09-817-910-7	Sequence 7, Appl
44	39	36.4	364	10	US-09-901-484A-5	Sequence 5, Appl
45	39	36.4	364	10	US-09-853-526-5	Sequence 5, Appl

ALIGNMENTS

RESULT 1

US-09-995-749A-10

Sequence 10, Application US/09995749A

Patent No. US2002015568A1

GENERAL INFORMATION:

APPLICANT: VAN GEEL-SCHUTTEN, GERRITDINA HENDRIKA

APPLICANT: DIJKHUIZEN, LOUBERT

APPLICANT: RAHAOUI, HAKIM

APPLICANT: LEER, ROBERT-JAN

TITLE OF INVENTION: NOVEL GLUCOSYLTRANSFERASES

FILE REFERENCE: B043388-CIP

CURRENT FILING DATE: 2001-11-29

PRIOR FILING DATE: 2000-06-28

PRIOR APPLICATION NUMBER: 09/604,957

PRIOR APPLICATION NUMBER: EPO 00201871.1

PRIOR FILING DATE: 2000-05-25

NUMBER OF SEQ ID NOS: 19

SOFTWARE: Patentin Ver. 2.1

SEQ ID NO 10

LENGTH: 545

TYPE: PRT

ORGANISM: Streptococcus mutans

US-09-995-749A-10

Query Match

Best Local Similarity 100.0%; Score 107; DB 10; Length 545;

Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MANYIFIRAHDSHVQTVIAKII 22

DB 154 MANYIFIRAHDSHVQTVIAKII 175

RESULT 2

US-09-740-274-6

Sequence 6, Application US/09740274

Patent No. US20020031826A1

```

; GENERAL INFORMATION:
; APPLICANT: Nichols, Scott E.
; TITLE OF INVENTION: Glucan-containing Compositions and Paper
; FILE REFERENCE: 0357CRD
; CURRENT APPLICATION NUMBER: US/09/740,274
; CURRENT FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: 09/210,361
; PRIOR FILING DATE: 1998-12-11
; PRIOR APPLICATION NUMBER: 09/007,999
; PRIOR FILING DATE: 1998-01-16
; PRIOR APPLICATION NUMBER: 08/478,704
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: 09/009,620
; PRIOR FILING DATE: 1998-01-20
; PRIOR APPLICATION NUMBER: 08/485,243
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: 09/008,172
; PRIOR FILING DATE: 1998-01-16
; PRIOR APPLICATION NUMBER: 08/482,711
; PRIOR FILING DATE: 1995-06-07
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 6
; LENGTH: 1430
; TYPE: PRT
; ORGANISM: streptococcus mutans
US-09-740-274-6
```

```

Query Match          100.0%; Score 107; DB 9; Length 1430;
Best Local Similarity 100.0%; Pred. No. 4,8e-09;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 MANYFIRAHDESEVQTVIAKII 22
Db      574 MANYFIRAHDESEVQTVIAKII 595
```

```

RESULT 3
US-09-995-749A-11
; Sequence 11, Application US/09995749A
; Patent No. US2002015568A1
; GENERAL INFORMATION:
; APPLICANT: VAN GEEL-SCHUTTEN, GERRITDINA HENDRIKA
; APPLICANT: DIJKHUIZEN, LIMBERT
; APPLICANT: RAHOUTI, HAKIM
; APPLICANT: LEER, ROBERT-VAN
; TITLE OF INVENTION: NOVEL GLUCOSYLTRANSFERASES
; FILE REFERENCE: BO43388-CIP
; CURRENT APPLICATION NUMBER: US/09/995,749A
; CURRENT FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: 09/604,957
; PRIOR FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: EPO 00201871.1
; PRIOR FILING DATE: 2000-05-25
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 11
; LENGTH: 522
; TYPE: PRT
; ORGANISM: Leuconostoc mesenteroides
US-09-995-749A-11
```

```

Query Match          80.4%; Score 86; DB 10; Length 522;
Best Local Similarity 80.0%; Pred. No. 4,9e-06;
Matches 16; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
```

```
QY      3 NYIFIRAHDESEVQTVIAKII 22
Db      148 NYIFIRAHDESEVQTVIAKII 167
```

```

RESULT 4
US-09-740-274-2
```

```

; Sequence 2, Application US/09740274
; Patent No. US20020031826A1
; GENERAL INFORMATION:
; APPLICANT: Nichols, Scott E.
; TITLE OF INVENTION: Glucan-containing Compositions and Paper
; FILE REFERENCE: 0357CRD
; CURRENT APPLICATION NUMBER: US/09/740,274
; CURRENT FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: 09/210,361
; PRIOR FILING DATE: 1998-12-11
; PRIOR APPLICATION NUMBER: 09/007,999
; PRIOR FILING DATE: 1998-01-16
; PRIOR APPLICATION NUMBER: 08/478,704
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: 09/009,620
; PRIOR FILING DATE: 1998-01-20
; PRIOR APPLICATION NUMBER: 08/485,243
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: 09/008,172
; PRIOR FILING DATE: 1998-01-16
; PRIOR APPLICATION NUMBER: 08/482,711
; PRIOR FILING DATE: 1995-06-07
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 1475
; TYPE: PRT
; ORGANISM: Streptococcus mutans
US-09-740-274-2
```

```

Query Match          67.3%; Score 72; DB 9; Length 1475;
Best Local Similarity 75.0%; Pred. No. 0.0035;
Matches 15; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
```

```
QY      3 NYIFIRAHDESEVQTVIAKII 22
Db      554 NYIFIRAHDESEVQTVIAKII 573
```

```

RESULT 5
US-09-740-274-4
; Sequence 4, Application US/09740274
; Patent No. US20020031826A1
; GENERAL INFORMATION:
; APPLICANT: Nichols, Scott E.
; TITLE OF INVENTION: Glucan-containing Compositions and Paper
; FILE REFERENCE: 0357CRD
; CURRENT APPLICATION NUMBER: US/09/740,274
; CURRENT FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: 09/210,361
; PRIOR FILING DATE: 1998-12-11
; PRIOR APPLICATION NUMBER: 09/007,999
; PRIOR FILING DATE: 1998-01-16
; PRIOR APPLICATION NUMBER: 08/478,704
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: 09/009,620
; PRIOR FILING DATE: 1998-01-20
; PRIOR APPLICATION NUMBER: 08/485,243
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: 09/008,172
; PRIOR FILING DATE: 1998-01-16
; PRIOR APPLICATION NUMBER: 08/482,711
; PRIOR FILING DATE: 1995-06-07
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 1375
; TYPE: PRT
; ORGANISM: streptococcus mutans
US-09-740-274-4
```

```

Query Match          63.6%; Score 68; DB 9; Length 1375;
Best Local Similarity 70.0%; Pred. No. 0.015;
```

Matches 14; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 3 NYIFRAHDESVQTVIAKII 22
Db 580 SYSFVRAHDNSQDQIQNAI 599

RESULT 6
US-09-995-749A-12
Sequence 12, Application US/09995749A
Patent No. US2002015568A1
GENERAL INFORMATION:
APPLICANT: VAN GEEL-SCHUTTEN, GERRITDINA HENDRIKA
APPLICANT: DIJKHUIZEN, LUBBERT
APPLICANT: RAHMOU, HAKIM
APPLICANT: LEER, ROBERT-JAN
TITLE OF INVENTION: NOVEL GLUCOSYLTRANSFERASES
FILE REFERENCE: B043388-CIP
CURRENT APPLICATION NUMBER: US/09/995,749A
CURRENT FILING DATE: 2001-11-29
PRIOR APPLICATION NUMBER: 09/604,957
PRIOR FILING DATE: 2000-06-28
PRIOR APPLICATION NUMBER: EPO 00201871.1
NUMBER OF SEQ ID NOS: 19
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 12
LENGTH: 584
TYPE: PRT
ORGANISM: Leuconostoc mesenteroides
US-09-995-749A-12

Query Match 49.5%; Score 53; DB 10; Length 584;
Best Local Similarity 55.6%; Pred. No. 1.8;
Matches 10; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 3 NYIFRAHDESVQTVIAK 20
Db 169 NYSFVRAHDYDADPIRK 186

RESULT 7
US-09-995-749A-13
Sequence 13, Application US/09995749A
Patent No. US2002015568A1
GENERAL INFORMATION:
APPLICANT: VAN GEEL-SCHUTTEN, GERRITDINA HENDRIKA
APPLICANT: DIJKHUIZEN, LUBBERT
APPLICANT: RAHMOU, HAKIM
APPLICANT: LEER, ROBERT-JAN
TITLE OF INVENTION: NOVEL GLUCOSYLTRANSFERASES
FILE REFERENCE: B043388-CIP
CURRENT APPLICATION NUMBER: US/09/995,749A
CURRENT FILING DATE: 2001-11-29
PRIOR APPLICATION NUMBER: 09/604,957
PRIOR FILING DATE: 2000-06-28
PRIOR APPLICATION NUMBER: EPO 00201871.1
NUMBER OF SEQ ID NOS: 19
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 13
LENGTH: 535
TYPE: PRT
ORGANISM: Lactobacillus reuteri
US-09-995-749A-13

Query Match 46.7%; Score 50; DB 10; Length 535;
Best Local Similarity 50.0%; Pred. No. 5.2;
Matches 10; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

QY 3 NYIFRAHDESVQTVIAKII 22
Db 146 NYSFVRAHDNSQDQIQNAI 165

RESULT 8
US-09-995-749A-2
Sequence 2, Application US/09995749A
Patent No. US2002015568A1
GENERAL INFORMATION:
APPLICANT: VAN GEEL-SCHUTTEN, GERRITDINA HENDRIKA
APPLICANT: DIJKHUIZEN, LUBBERT
APPLICANT: RAHMOU, HAKIM
APPLICANT: LEER, ROBERT-JAN
TITLE OF INVENTION: NOVEL GLUCOSYLTRANSFERASES
FILE REFERENCE: B043388-CIP
CURRENT APPLICATION NUMBER: US/09/995,749A
CURRENT FILING DATE: 2001-11-29
PRIOR APPLICATION NUMBER: 09/604,957
PRIOR FILING DATE: 2000-06-28
PRIOR APPLICATION NUMBER: EPO 00201871.1
NUMBER OF SEQ ID NOS: 19
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 2
LENGTH: 1781
TYPE: PRT
ORGANISM: Lactobacillus reuteri
US-09-995-749A-2

Query Match 46.7%; Score 50; DB 10; Length 1781;
Best Local Similarity 50.0%; Pred. No. 21;
Matches 10; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

QY 3 NYIFRAHDESVQTVIAKII 22
Db 1125 NYSFVRAHDNSQDQIQNAI 1144

RESULT 9
US-09-758-269-2
Sequence 2, Application US/09758269
Patent No. US20020104120A1
GENERAL INFORMATION:
APPLICANT: IUCHI, SATOSHI
APPLICANT: KOBAYASHI, MASATOMO
APPLICANT: SHINOZAKI, KAZUO
TITLE OF INVENTION: TRANSGENE PLANTS CARRYING NEOXANTHIN
FILE REFERENCE: 3914-3
CURRENT APPLICATION NUMBER: US/09/758,269
CURRENT FILING DATE: 2001-01-12
PRIOR APPLICATION NUMBER: JP 2001-003476
PRIOR FILING DATE: 2001-01-11
PRIOR APPLICATION NUMBER: JP 2000-010056
PRIOR FILING DATE: 2000-01-13
NUMBER OF SEQ ID NOS: 33
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 2
LENGTH: 583
TYPE: PRT
ORGANISM: Arabidopsis thaliana
US-09-758-269-2

Query Match 41.1%; Score 44; DB 10; Length 583;
Best Local Similarity 47.4%; Pred. No. 58;
Matches 9; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

QY 4 YIFRAHDESVQTVIAKII 22
Db 530 YIFCHVADDETKTSEJQII 548

RESULT 10
US-10-032-585-7827
Sequence 7827, Application US/10032585

```
Publication No. US20030180953A1
GENERAL INFORMATION:
APPLICANT: Terry, Roemer D.
APPLICANT: Bo, Jians
APPLICANT: Charles, Boone
APPLICANT: Howard, Busey
TITLE OF INVENTION: Gene Disruption Methodologies for Drug Target Discovery
FILE REFERENCE: 10182-005-999
CURRENT APPLICATION NUMBER: US/10/032,585
CURRENT FILING DATE: 2001-12-20
NUMBER OF SEQ ID NOS: 8000
SOFTWARE: PatentIn version 3.1
SEQ ID NO 7827
LENGTH: 327
TYPE: PRT
ORGANISM: Candida albicans
US-10-032-585-7827
```

```
Query Match          39.3%  Score 42; DB 12; Length 327;
Best Local Similarity 44.4%  Pred. No. 64;
Matches 8; Conservative 3; Mismatches 7; Indels 0; Gaps 0;
```

```
QY 5 IFIRAHDESVQTVIAKII 22
DB 290 IFVDGEDEDHYHYIEKIV 307
```

```
RESULT 11
US-10-007-280A-214
Sequence 57, Application US/10007280A
Publication No. US20030059784A1
GENERAL INFORMATION:
APPLICANT: Sun, Yongming
APPLICANT: Recipon, Heve
APPLICANT: Salceda, Susana
APPLICANT: Chenchua, Liu
TITLE OF INVENTION: Compositions and Methods Relating to Ovary Specific Genes and Pro
FILE REFERENCE: DEX-0257
CURRENT APPLICATION NUMBER: US/10/007,280A
CURRENT FILING DATE: 2001-11-07
PRIOR APPLICATION NUMBER: US 60/246,640
PRIOR FILING DATE: 2000-11-08
NUMBER OF SEQ ID NOS: 238
SOFTWARE: PatentIn version 3.1
SEQ ID NO 214
LENGTH: 67
TYPE: PRT
ORGANISM: Homo sapien
US-10-007-280A-214
```

```
Query Match          37.4%  Score 40; DB 15; Length 67;
Best Local Similarity 40.0%  Pred. No. 22;
Matches 6; Conservative 4; Mismatches 5; Indels 0; Gaps 0;
```

```
QY 6 FIRAHDESVQTVIAK 20
DB 49 FIVAHNOEINLILTK 63
```

```
RESULT 12
US-10-270-875-57
Sequence 57, Application US/10270875
Publication No. US20030082741A1
GENERAL INFORMATION:
APPLICANT: Sigridur Hjorleifsdottir
APPLICANT: Gudmundur O. Hregvaldsson
APPLICANT: Olafur H. Fridjonsson
APPLICANT: Arnthor Aeyarsson
APPLICANT: Jakob K. Kristjansson
TITLE OF INVENTION: Bacteriophage RM378 of a Thermophilic
FILE REFERENCE: 2739.1001-001
CURRENT APPLICATION NUMBER: US/10/270,875
```

```
CURRENT FILING DATE: 2002-10-11
PRIOR APPLICATION NUMBER: US/09/585,858
PRIOR FILING DATE: 2000-12-18
PRIOR APPLICATION NUMBER: 60/137,120
PRIOR FILING DATE: 1999-06-02
NUMBER OF SEQ ID NOS: 73
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 57
LENGTH: 349
TYPE: PRT
ORGANISM: Bacteriophage RM378
US-10-270-875-57
```

```
Query Match          37.4%  Score 40; DB 15; Length 349;
Best Local Similarity 36.8%  Pred. No. 1.5e+02;
Matches 7; Conservative 5; Mismatches 7; Indels 0; Gaps 0;
```

```
QY 4 YIFIRAHDESVQTVIAKII 22
DB 175 YLIRSHIEVNVISDKLL 193
```

```
RESULT 13
US-10-270-878-57
Sequence 57, Application US/10270878
Publication No. US20030082790A1
GENERAL INFORMATION:
APPLICANT: Sigridur Hjorleifsdottir
APPLICANT: Gudmundur O. Hregvaldsson
APPLICANT: Olafur H. Fridjonsson
APPLICANT: Arnthor Aeyarsson
APPLICANT: Jakob K. Kristjansson
TITLE OF INVENTION: Bacteriophage RM378 of a Thermophilic
FILE REFERENCE: 2739.1001-001
CURRENT APPLICATION NUMBER: US/10/270,878
CURRENT FILING DATE: 2002-10-11
PRIOR APPLICATION NUMBER: US/09/585,858
PRIOR FILING DATE: 2000-12-18
NUMBER OF SEQ ID NOS: 73
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 57
LENGTH: 349
TYPE: PRT
ORGANISM: Bacteriophage RM378
US-10-270-878-57
```

```
Query Match          37.4%  Score 40; DB 15; Length 349;
Best Local Similarity 36.8%  Pred. No. 1.5e+02;
Matches 7; Conservative 5; Mismatches 7; Indels 0; Gaps 0;
```

```
QY 4 YIFIRAHDESVQTVIAKII 22
DB 175 YLIRSHIEVNVISDKLL 193
```

```
RESULT 14
US-10-270-786-57
Sequence 57, Application US/10270786
Publication No. US20030087392A1
GENERAL INFORMATION:
APPLICANT: Sigridur Hjorleifsdottir
APPLICANT: Gudmundur O. Hregvaldsson
APPLICANT: Olafur H. Fridjonsson
APPLICANT: Arnthor Aeyarsson
APPLICANT: Jakob K. Kristjansson
TITLE OF INVENTION: Bacteriophage RM378 of a Thermophilic
FILE REFERENCE: 2739.1001-001
CURRENT APPLICATION NUMBER: US/10/270,786
CURRENT FILING DATE: 2002-10-11
PRIOR APPLICATION NUMBER: US/09/585,858
PRIOR FILING DATE: 2000-12-18
```

Query Match	37.4%;	Score 40;	DB 15;	Length 349;
Best Local Similarity	36.8%;	Pred. No. 1.5e+02;		
Matches	7;	Conservative	5;	Mismatches 7;
				Indels

QY 4 YIFIRAHSEVQGVIAKTI 22
| : | : | : | : | : | :
Db 175 YLIRSHIEVNIVISDKL 193

```

RESULT 15
US-10-270-710-57
? Sequence 57, Application US/10270710
? Publication No. US20030092128A1
? GENERAL INFORMATION:
? APPLICANT: Sigridur Hjorleifsdottir
? APPLICANT: Gudmundur O. Hreggvidsson
? APPLICANT: Olafur H. Fridjonsson
? APPLICANT: Arnthor Aevirsson
? APPLICANT: Jakob K. Kristjansson
? TITLE OF INVENTION: Bacteriophage RM378 of a Thermophilic
? TITLE OF INVENTION: Host Organism
? FILE REFERENCE: 2739, 1001-001
? CURRENT APPLICATION NUMBER: US/10/270,710
? CURRENT FILING DATE: 2002-10-11
? PRIOR APPLICATION NUMBER: 2002-10-11
? PRIOR FILING DATE: 2000-12-18
? PRIOR APPLICATION NUMBER: 60/137,120
? PRIOR FILING DATE: 1999-06-02
? NUMBER OF SEQ ID NOS: 73
? SOFTWARE: FastSeq for Windows Version 4.0
? SEQ ID NO 57
? LENGTH: 349
? TYPE: PRF
? ORGANISM: Bacteriophage RM378
? US-10-270-710-57

```

Query March	37.4%	Score 40;	DB 15;	Length 349;
Best local	Similarity 36.8%;	Pred. NO. 1.5e+02;		
Matches	7;	Conservative	5;	Mismatches 7;
				Indels 0;
				Gaps 0;
Qy	4	YIFRADSEYQVIYAKII	22	
	:	:	:	:
	:	:	:	:
	:	:	:	:
Db	175	YLIRSHIEVNVISDKIL	193	

Search completed: November 13, 2003, 10:29:05
Job time : 29.7773 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 13, 2003, 09:31:40 ; Search time 40.3507 Seconds
(without alignments)
86.541 Million cell updates/sec

Title: US-09-290-049A-16

Perfect score: 107

Sequence: 1 MANYPIRAHDSVQTVIAKII 22

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_19Jun03.*
1: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*
24: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	107	100.0	1430	23	AAU98029	S. mutans glucosyl
2	107	100.0	1430	23	AAU98043	S. mutans glucosyl
3	101	94.4	1430	23	AAU98041	S. mutans glucosyl
4	101	94.4	1430	23	AAU98042	S. mutans glucosyl
5	101	94.4	1430	23	AAU98044	S. mutans glucosyl
6	101	94.4	1430	23	AAU98045	S. mutans glucosyl
7	89	83.2	1577	17	AAU91047	Alpha-D-glucosyltr
8	89	83.2	2835	23	ABB98574	Dextran saccharase
9	86	80.4	1527	23	AAU80055	Leuconostoc mesent

10	84	78.5	1475	23	AAU98036
11	84	78.5	1475	23	AAU98037
12	84	78.5	1475	23	AAU98038
13	78	72.9	1475	23	AAU98032
14	78	72.9	1475	23	AAU98034
15	78	72.9	1475	23	AAU98035
16	72	67.3	1475	23	AAU98027
17	72	67.3	1475	23	AAU98030
18	72	67.3	1475	23	AAU98031
19	72	67.3	1475	23	AAU98033
20	72	67.3	1475	23	AAU98039
21	72	67.3	1475	23	AAU98040
22	70	65.4	15	23	ABB98579
23	70	65.4	15	23	ABB98577
24	68	63.6	15	23	ABB98653
25	68	63.6	15	23	ABB98654
26	68	63.6	15	23	ABB98655
27	68	63.6	1375	23	AAU98028
28	68	63.6	1375	23	AAU98028
29	67	62.6	15	23	ABB98652
30	67	62.6	1017	23	AAU98028
31	67	62.6	1476	23	AAU98028
32	64	59.8	1592	14	AAU98028
33	60	56.1	15	23	ABB98650
34	55	51.4	15	23	ABB98651
35	53	49.5	2057	21	ABB10667
36	50	46.7	1781	23	AAU98028
37	47	43.9	474	22	ABB64936
38	46.5	43.5	866	22	ABB64833
39	46	43.0	881	23	ABB65773
40	45	42.1	437	23	ABB29490
41	45	42.1	655	22	ABB59162
42	44	41.1	282	21	AAU98031
43	44	41.1	336	21	AAU98031
44	44	41.1	423	21	AAU98031
45	44	41.1	425	21	AAU98031

ALIGNMENTS

RESULT 1	
AAU98029	
ID	AAU98029 standard; Protein; 1430 AA.
XX	
AC	AAU98029;
XX	
DT	27-AUG-2002 (first entry)
XX	
DE	S. mutans glucosyltransferase GTFD.
XX	
KW	Glucosyltransferase; GTFD; transgenic plant; paper siring;
KW	coating composition; glucan; starch; latex; thermoplastic molecule;
KW	amyloplast; vacuole; paper manufacture.
XX	
OS	Streptococcus mutans.
XX	
FN	US2002031826-A1.
XX	
PD	14-MAR-2002.
XX	
PF	19-DEC-2000; 2000US-0740274.
XX	
ER	11-DEC-1998; 98US-0210361.
ER	07-JUN-1995; 95US-0478704.
ER	07-JUN-1995; 95US-0482711.
PR	07-JUN-1995; 95US-0485243.
PR	16-JAN-1998; 98US-0007999.
PR	16-JAN-1998; 98US-0008172.
PR	20-JAN-1998; 98US-0009620.
XX	
PA	(NICH//) NICHOLS S E.
XX	

Pt	NicholsSE;
Xx	
Dr	WPI; 2002-414332/44.
Dn	N-PBDB; ABRK52940.
Xx	
Pt	Glucosyltransferase B or D protein useful for producing a glucan useful
Pt	as substitutes for and additions to modified starch and latexes in
Pt	paper manufacture, comprises mutations in specific positions -
Xx	
P5	Disclosure; Page 38-42; 44pp; English.
Xx	
Cc	The invention an isolated protein comprising a glucosyltransferase
Cc	(GTF) polypeptide having changes at position from I448V, D457N,
Cc	D567T, K1014T, D457N/D567T, D457N/D571K, D567T/D571K,
Cc	D567T/D571K/K1014T, I448V/D457N/D567T/D571K/K779Q/K1014T,
Cc	V156N/Y170A/I171A, and K779Q or a GTF D polypeptide having
Cc	changes at positions from T589D, T589E, N471D, N471D/T589D, and
Cc	N471D/T589E. Also included are a glucan produced by the GTF mutant,
Cc	an isolated polynucleotide which encodes P1 or P2, or its complementary
Cc	polynucleotide, a ribonucleic acid sequence encoding the GTF mutant,
Cc	an expression cassette comprising the polynucleotide operably linked to a
Cc	promoter, a vector comprising the expression cassette, host cell
Cc	introduced with the vector, a transgenic plant comprising the
Cc	vector, a seed or tuber from the transgenic plant, a paper sizing and/or
Cc	coating composition comprising a glucan produced in a plant transformed
Cc	with a gene encoding the mutant GTF, wild type or, starch, a latex,
Cc	thermoplastic molecule or their combinations or glucan and starch where
Cc	the glucan is produced in the amyloplast and/or vacuole or a maize line
Cc	deficient in starch biosynthesis, transformed with a gene encoding a
Cc	glucosyltransferase B or D enzyme, wild-type or mutant and a paper
Cc	comprising the glucan (paper sizing/coating agent). The vector is useful
Cc	for producing a glucan in a plant. The method comprises transforming a
Cc	plant cell with the vector, growing the plant cell under plant growing
Cc	conditions to produce a regenerated plant and inducing expression of the
Cc	polynucleotide for a time sufficient to produce the glucan in the
Cc	regenerated plant, where the vector contains a transit sequence from
Cc	ribulose biphosphate carboxylase small subunit, waxy, brittle-1 and
Cc	chlorophyll AB binding protein to produce a transgenic plant, and glucan
Cc	is produced in the amyloplast of potato or the vacuole of sugar beet.
Cc	Glucans are useful as substitutes for and additions to modified starch
Cc	and latexes in paper manufacture. Unlike prior art techniques, which
Cc	require input materials that produce chemical effluents, paper
Cc	manufacture utilizing the glucan produced by GTF, which utilizes
Cc	biologically produced input materials, is more cost-effective and
Cc	environmentally friendly. Moreover, glucans also exhibit thermoplastic
Cc	properties and impart gloss to the paper during coating step.
Cc	The present sequence represents GTFD.
Xx	
SQ	Sequence 1430 AA;
	Query Match 100.0%; Score 107; DB 23; Length 1430;
	Best Local Similarity 100.0%; Pred. No. 4,7e-10;
	Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy	1 MANYIFIRAHDSVEQTIVAKII 22
Dd	574 MANYIFIRAHDSVEQTIVAKII 595
RESULT 2	
ID	AAU98043 standard; Protein; 1430 AA.
XX	
AC	AAU98043;
XX	
DT	27-AUG-2002 (first entry)
XX	
DE	S. mutans glucosyltransferase GTFD mutant N471D.
Kw	Glucosyltransferase; GTFD; transgenic plant; paper sizing;
Kw	coating composition; glucan; starch; latex; thermoplastic molecule;
Kw	amyloplast; vacuole; paper manufacture; mutant; muten.
XX	

OS	Streptococcus mutans.
XX	Synthetic.
PH	Key
FT	Misc-difference 471 Location/Qualifiers
XX	/note= "Wild-type Asn substituted by Asp"
PN	US2002031826-A1.
XX	
PD	14-MAR-2002.
XX	
PF	19-DEC-2000; 2000US-0740274.
XX	
PR	11-DEC-1998; 98US-0210361.
PR	07-JUN-1995; 95US-0478704.
PR	07-JUN-1995; 95US-0482711.
PR	07-JUN-1995; 95US-0485243.
PR	16-JAN-1998; 98US-0007999.
PR	16-JAN-1998; 98US-0008172.
XX	20-JAN-1998; 98US-0009620.
PA	(NICH/) NICHOLS S E.
XX	
P1	Nichols SE;
XX	
DR	WPI; 2002-414332/44.
XX	
PT	Glucosyltransferase B or D protein useful for producing a glucan useful
PT	as substitutes for and additions to modified starch and latexes in
PT	paper manufacture, comprises mutations in specific positions -
XX	
PS	Claim 36; Page -: 44pp; English.
CC	The invention an isolated protein comprising a glucosyltransferase
CC	(GTF) B polypeptide having changes at position from I448V, D457N,
CC	D567T, K1014E, D457N/D567T, D457N/D571K, D567T/D571K,
CC	D567T/D571K/K1014I, I448V/D457N/D567T/D571K/K779Q/K1014T,
CC	Y169A/Y101A/Y101A, and K779Q or a GTF D polypeptide having
CC	changes at positions from T589D, T589E, N471D, N471D/T589D, and
CC	N471D/T589E. Also included are a glucan produced by the GTF mutant,
CC	an isolated polynucleotide which encodes P1 or P2, or its complementary
CC	polynucleotide, a ribonucleic acid sequence encoding the GTF mutant,
CC	an expression cassette comprising the polynucleotide operably linked to
CC	promoter, a vector comprising the expression cassette, host cell
CC	introduced with the vector, a transgenic plant comprising the
CC	vector, a seed or tuber from the transgenic plant, a paper sizing and/or
CC	coating composition comprising a glucan produced in a plant transformed
CC	with a gene encoding the mutant GTF, wild type or, starch, a latex,
CC	thermoplastic molecule or their combinations or glucan and starch where
CC	the glucan is produced in the amyloplast and/or vacuole or a maize line
CC	deficient in starch biosynthesis, transformed with a gene encoding a
CC	glucosyltransferase B or D enzyme, wild-type or mutant and a paper
CC	comprising the glucan (paper sizing/coating agent). The vector is useful
CC	for producing a glucan in a plant. The method comprises transforming a
CC	plant cell with the vector, growing the plant cell under plant growing
CC	conditions to produce a regenerated plant and inducing expression of the
CC	polynucleotide for a time sufficient to produce the glucan in the
CC	regenerated plant, where the vector contains a transit sequence from
CC	ribulose biphosphate carboxylase small subunit, waxy, brittle-1 and
CC	chlorophyll AB binding protein to produce a transgenic plant, and glucan
CC	is produced in the amyloplast of potato or the vacuole of sugar beet.
CC	Glucans are useful as substitutes for and additions to modified starch
CC	and latexes in paper manufacture. Unlike prior art techniques, which
CC	require input materials that produce chemical effluents, paper
CC	manufacture utilizing the glucan produced by GTF, which utilizes
CC	biologically produced input materials, is more cost-effective and
CC	environmentally friendly. Moreover, glucans also exhibit thermoplastic
CC	properties and impart gloss to the paper during coating step.
CC	The present sequence represents a GTF mutant of the invention but
CC	Note: The present sequence is not shown in the specification but
CC	was created by the indexer using the GTF sequence appearing as AAU98029
CC	and the information in claim 36.

SQ Sequence 1430 AA;
 Query Match 100.0%; Score 107; DB 23; Length 1430;
 Best Local Similarity 100.0%; Pred. No. 4.7e-10;
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MANYIFIRAHSEVQVIKII 22
 DB 574 MANYIFIRAHSEVQVIKII 595
 RESULT 3
 AAU98041 ID AAU98041 standard; Protein; 1430 AA.
 AC AAU98041;
 DT 27-AUG-2002 (first entry)
 DE S. mutans glucosyltransferase GTFD mutant T589D.
 KM Glucosyltransferase; GTFD; transgenic plant; paper sizing;
 KW coating composition; glucan; starch; latex; thermoplastic molecule;
 KM amyloplast; vacuole; paper manufacture; mutant; muteln.
 OS Streptococcus mutans.
 OS Synthetic.
 FH Key Location/Qualifiers
 FT Misc-difference 589 /note= "Wild-type Thr substituted by Asp"
 FT US2002031826-A1.
 PD 14-MAR-2002.
 PF 19-DEC-2000; 2000US-0740274.
 PR 11-DEC-1998; 98US-0210361.
 PR 07-JUN-1995; 95US-0478704.
 PR 07-JUN-1995; 95US-0482711.
 PR 07-JUN-1995; 95US-0485243.
 PR 16-JAN-1998; 98US-0007999.
 PR 16-JAN-1998; 98US-0008172.
 PR 20-JAN-1998; 98US-0009620.
 PA (NICH/) NICHOLS S E.
 PI Nichols SE;
 DR WPI; 2002-414332/44.
 PT Glucosyltransferase B or D protein useful for producing a glucan useful
 PT as substitutes for and additions to modified starch and latexes in
 PT paper manufacture, comprises mutations in specific positions -
 XX Claim 36; Page -; 44pp; English.
 CC The invention an isolated protein comprising a glucosyltransferase
 CC (GTF) B polypeptide having changes at position from 1448V, D457N,
 CC D567T, K1014T, D457N/D567T, D457N/D571K, D567T/D571K,
 CC D567T/D571K/K1014T, 1448V/D457N/D567T/D571K/K790/K1014T,
 CC Y169A/Y170A/Y171A, and K779Q or a GTF D polypeptide having
 CC changes at positions from T589D, T589E, N471D, N471D/T589D and
 CC N471D/T589E. Also included are a glucan produced by the GTF mutant,
 CC an isolated polynucleotide which encodes P1 or P2, or its complementary
 CC polynucleotide, a ribonucleic acid sequence encoding the GTF mutant,
 CC an expression cassette comprising the polynucleotide operably linked to a
 CC promoter, a vector comprising the expression cassette, host cell
 CC introduced with the vector, a transgenic plant comprising the
 CC vector, a seed or tuber from the transgenic plant, a paper sizing and/or
 CC coating composition comprising a glucan produced in a plant transformed
 CC with a gene encoding the mutant GTF, wild type or, starch, a latex,
 CC

CC thermoplastic molecule or their combinations or glucan and starch where
 CC the glucan is produced in the amyloplast and/or vacuole or a maize line
 CC deficient in starch biosynthesis, transformed with a gene encoding a
 CC glucosyltransferase B or D enzyme, wild-type or mutant and a paper
 CC comprising the glucan (paper sizing/coating agent). The vector is useful
 CC for producing a glucan in a plant. The method comprises transforming a
 CC plant cell with the vector, growing the plant cell under plant growing
 CC conditions to produce a regenerated plant and inducing expression of the
 CC polynucleotide for a time sufficient to produce the glucan in the
 CC regenerated plant, where the vector contains a transit sequence from
 CC ribulose biphosphate carboxylase small subunit, waxy, brittle-1 and
 CC chlorophyll AB binding protein to produce a transgenic plant, and glucan
 CC is produced in the amyloplast of potato or the vacuole of sugar beet.
 CC Glucans are useful as substitutes for and additions to modified starch
 CC and latexes in paper manufacture. Unlike prior art techniques, which
 CC require input materials that produce chemical effluents, paper
 CC manufacture utilizing the glucan produced by GTF, which utilizes
 CC biologically produced input materials, is more cost-effective and
 CC environmentally friendly. Moreover, glucans also exhibit thermoplastic
 CC properties and impart gloss to the paper during coating step.
 CC The present sequence represents a GTFD mutant of the invention.
 CC Note: The present sequence is not shown in the specification but
 CC was created by the indexer using the GTFD sequence appearing as AAU98029
 CC and the information in claim 36.
 SQ Sequence 1430 AA;
 Query Match 94.4%; Score 101; DB 23; Length 1430;
 Best Local Similarity 95.5%; Pred. No. 5.7e-09;
 Matches 21; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 MANYIFIRAHSEVQVIKII 22
 DB 574 MANYIFIRAHSEVQVIKII 595
 RESULT 4
 AAU98042 ID AAU98042 standard; Protein; 1430 AA.
 AC AAU98042;
 DT 27-AUG-2002 (first entry)
 DE S. mutans glucosyltransferase GTFD mutant T589E.
 KM Glucosyltransferase; GTFD; transgenic plant; paper sizing;
 KW coating composition; glucan; starch; latex; thermoplastic molecule;
 KM amyloplast; vacuole; paper manufacture; mutant; muteln.
 OS Streptococcus mutans.
 OS Synthetic.
 FH Key Location/Qualifiers
 FT Misc-difference 589 /note= "Wild-type Thr substituted by Glu"
 FT US2002031826-A1.
 PD 14-MAR-2002.
 PF 19-DEC-2000; 2000US-0740274.
 PR 11-DEC-1998; 98US-0210361.
 PR 07-JUN-1995; 95US-0478704.
 PR 07-JUN-1995; 95US-0482711.
 PR 07-JUN-1995; 95US-0485243.
 PR 16-JAN-1998; 98US-0007999.
 PR 16-JAN-1998; 98US-0008172.
 PR 20-JAN-1998; 98US-0009620.
 PA (NICH/) NICHOLS S E.
 PI Nichols SE;
 DR WPI; 2002-414332/44.
 PT Glucosyltransferase B or D protein useful for producing a glucan useful
 PT as substitutes for and additions to modified starch and latexes in
 PT paper manufacture, comprises mutations in specific positions -
 XX Claim 36; Page -; 44pp; English.
 CC The invention an isolated protein comprising a glucosyltransferase
 CC (GTF) B polypeptide having changes at position from 1448V, D457N,
 CC D567T, K1014T, D457N/D567T, D457N/D571K, D567T/D571K,
 CC D567T/D571K/K1014T, 1448V/D457N/D567T/D571K/K790/K1014T,
 CC Y169A/Y170A/Y171A, and K779Q or a GTF D polypeptide having
 CC changes at positions from T589D, T589E, N471D, N471D/T589D and
 CC N471D/T589E. Also included are a glucan produced by the GTF mutant,
 CC an isolated polynucleotide which encodes P1 or P2, or its complementary
 CC polynucleotide, a ribonucleic acid sequence encoding the GTF mutant,
 CC an expression cassette comprising the polynucleotide operably linked to a
 CC promoter, a vector comprising the expression cassette, host cell
 CC introduced with the vector, a transgenic plant comprising the
 CC vector, a seed or tuber from the transgenic plant, a paper sizing and/or
 CC coating composition comprising a glucan produced in a plant transformed
 CC with a gene encoding the mutant GTF, wild type or, starch, a latex,
 CC

PI Nichols SE;
 DR WPI: 2002-414332/44.
 XX
 XX Glucosyltransferase B or D protein useful for producing a glucan useful
 PT as substitutes for and additions to modified starch and latexes in
 PT paper manufacture, comprises mutations in specific positions -
 XX
 XX Claim 36; Page -: 44pp; English.
 XX
 CC The invention an isolated protein comprising a glucosyltransferase
 CC (GTF) B polypeptide having changes at position from 1448V, D457N,
 CC D567T, K1014T, D457N/D567T, D457N/D571K, D567T/D571K,
 CC Y169A/Y170A/K1014T, 1448V/D457N/D567T/D571K/K779Q/K1014T,
 CC K169A/Y170A/Y171A, and K779Q or a GTF D polypeptide having
 CC changes at positions from T589D, T589E, N471D, N471D/T589D, and
 CC N471D/T589E. Also included are a glucan produced by the GTF mutant,
 CC an isolated polynucleotide which encodes P1 or P2, or its complementary
 CC polynucleotide, a ribonucleic acid sequence encoding the GTF mutant,
 CC an expression cassette comprising the polynucleotide operably linked to a
 CC promoter, a vector comprising the expression cassette, host cell
 CC introduced with the vector, a transgenic plant comprising the
 CC vector, a seed or tuber from the transgenic plant, a paper sizing and/or
 CC coating composition comprising a glucan produced in a plant transformed
 CC with a gene encoding the mutant GTF, wild type or, starch, a latex,
 CC thermoplastic molecule or their combinations or glucan and starch where
 CC the glucan is produced in the amyloplast and/or vacuole or a maize line
 CC deficient in starch biosynthesis, transformed with a gene encoding a
 CC glucosyltransferase B or D enzyme, wild-type or mutant and a paper
 CC comprising the glucan (paper sizing/coating agent). The vector is useful
 CC for producing a glucan in a plant. The method comprises transforming a
 CC plant cell with the vector, growing the plant cell under plant growing
 CC conditions to produce a regenerated plant and inducing expression of the
 CC polynucleotide for a time sufficient to produce the glucan in the
 CC regenerated plant, where the vector contains a transit sequence from
 CC chlorophyll AB binding protein to produce a transgenic plant, and glucan
 CC is produced in the amyloplast of potato or the vacuole of sugar beet.
 CC Glucans are useful as substitutes for and additions to modified starch
 CC and latexes in paper manufacture. Unlike prior art techniques, which
 CC require input materials that produce chemical effluents, paper
 CC manufacture utilizing the glucan produced by GTF, which utilizes
 CC biologically produced input materials, is more cost-effective and
 CC environmentally friendly. Moreover, glucans also exhibit thermoplastic
 CC properties and impart gloss to the paper during coating step.
 CC The present sequence represents a GTF mutant of the invention.
 CC Note: The present sequence is not shown in the specification but
 CC was created by the indexer using the GTF sequence appearing as AAU98029
 CC and the information in claim 36.
 CC
 XX
 SQ Sequence 1430 AA;
 Query Match 94.4%; Score 101; DB 23; Length 1430;
 Best Local Similarity 95.5%; Pred. No. 5; 7e-09;
 Matches 21; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 Oy 1 MANYIFIRAHDSVQVIAKII 22
 DB 574 MANYIFIRAHDSVQVIAKII 595
 RESULT 5
 AAU98044
 ID AAU98044 standard; Protein; 1430 AA.
 XX
 AC AAU98044;
 XX
 DT 27-AUG-2002 (first entry)
 XX
 XX S. mutans glucosyltransferase GTFD mutant N471D/T589D.
 DE
 XX Glucosyltransferase; GTFD; transgenic plant; paper sizing;
 KM coating composition; glucan; starch; latex; thermoplastic molecule;
 CC

KM amyloplast; vacuole; paper manufacture; mutant; mutain.
 XX
 XX Streptococcus mutans.
 OS Synthetic.
 OS
 XX
 XX Key Location/Qualifiers
 FT Misc-difference 471
 FT Misc-difference 589 /note= "Wild-type Asn substituted by Asp"
 FT Misc-difference 589 /note= "Wild-type Thr substituted by Asp"
 XX
 XX US2002031826-A1.
 PN
 XX
 XX 14-MAR-2002.
 PD
 XX
 XX 19-DEC-2000; 2000US-0740274.
 PF
 XX
 XX 11-DEC-1998; 98US-0210361.
 PR 07-JUN-1995; 95US-0478704.
 PR 07-JUN-1995; 95US-0482711.
 PR 07-JUN-1995; 95US-0485243.
 PR 16-JUN-1998; 98US-0007999.
 PR 16-JUN-1998; 98US-0008172.
 PR 20-JUN-1998; 98US-0009620.
 XX
 XX (NICH/) NICHOLS S E.
 PA
 PI Nichols SE;
 XX
 XX WPI: 2002-414332/44.
 XX
 XX Glucosyltransferase B or D protein useful for producing a glucan useful
 PT as substitutes for and additions to modified starch and latexes in
 PT paper manufacture, comprises mutations in specific positions -
 XX
 XX Claim 36; Page -: 44pp; English.
 XX
 CC The invention an isolated protein comprising a glucosyltransferase
 CC (GTF) B polypeptide having changes at position from 1448V, D457N,
 CC D567T, K1014T, D457N/D567T, D457N/D571K, D567T/D571K,
 CC Y169A/Y170A/K1014T, 1448V/D457N/D567T/D571K/K779Q/K1014T,
 CC K169A/Y170A/Y171A, and K779Q or a GTF D polypeptide having
 CC changes at positions from T589D, T589E, N471D, N471D/T589D, and
 CC N471D/T589E. Also included are a glucan produced by the GTF mutant,
 CC an isolated polynucleotide which encodes P1 or P2, or its complementary
 CC polynucleotide, a ribonucleic acid sequence encoding the GTF mutant,
 CC an expression cassette comprising the polynucleotide operably linked to a
 CC promoter, a vector comprising the expression cassette, host cell
 CC introduced with the vector, a transgenic plant comprising the
 CC vector, a seed or tuber from the transgenic plant, a paper sizing and/or
 CC coating composition comprising a glucan produced in a plant transformed
 CC with a gene encoding the mutant GTF, wild type or, starch, a latex,
 CC thermoplastic molecule or their combinations or glucan and starch where
 CC the glucan is produced in the amyloplast and/or vacuole or a maize line
 CC deficient in starch biosynthesis, transformed with a gene encoding a
 CC glucosyltransferase B or D enzyme, wild-type or mutant and a paper
 CC comprising the glucan (paper sizing/coating agent). The vector is useful
 CC for producing a glucan in a plant. The method comprises transforming a
 CC plant cell with the vector, growing the plant cell under plant growing
 CC conditions to produce a regenerated plant and inducing expression of the
 CC polynucleotide for a time sufficient to produce the glucan in the
 CC regenerated plant, where the vector contains a transit sequence from
 CC ribulose biphosphate carboxylase small subunit, waxy, brittle-1 and
 CC chlorophyll AB binding protein to produce a transgenic plant, and glucan
 CC is produced in the amyloplast of potato or the vacuole of sugar beet.
 CC Glucans are useful as substitutes for and additions to modified starch
 CC and latexes in paper manufacture. Unlike prior art techniques, which
 CC require input materials that produce chemical effluents, paper
 CC manufacture utilizing the glucan produced by GTF, which utilizes
 CC biologically produced input materials, is more cost-effective and
 CC environmentally friendly. Moreover, glucans also exhibit thermoplastic
 CC properties and impart gloss to the paper during coating step.
 CC The present sequence represents a GTFD mutant of the invention.
 CC

CC Note: The present sequence is not shown in the specification but
CC was created by the indexer using the GTFD sequence appearing as AAU98029
CC and the information in claim 36.

CC Sequence 1430 AA;

Query Match 94.4%; Score 101; DB 23; Length 1430;
Best Local Similarity 95.5%; Pred. No. 5.7e-09;
Matches 21; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Db 1 MANYIFIRAHDSVQVYIAKII 22
574 MANYIFIRAHDSVQVYIAKII 595

RESULT 6
AAU98045
ID AAU98045 standard; Protein; 1430 AA.

AC AAU98045;
XX
DT 27-AUG-2002 (first entry)
XX
DE S. mutans glucosyltransferase GTFD mutant N471D/T589E.
XX
KW Glucosyltransferase; GTFD; transgenic plant; paper sizing;
KW coating composition; glucan; starch; latex; thermoplastic molecule;
KW amyloplast; vacuole; paper manufacture; mutant; mucelin.
XX
OS Streptococcus mutans.
XX
OS Synthetic.

Key Location/Qualifiers
FT Misc-difference 471 /note= "Wild-type Asn substituted by Asp"
FT Misc-difference 589 /note= "Wild-type Thr substituted by Glu"
XX
PN US2002031826-A1.

PD 14-MAR-2002.
XX
PF 19-DEC-2000; 2000US-0740274.

XX
PR 11-DEC-1998; 98US-0210361.
XX
PR 07-JUN-1995; 95US-0478704.
XX
PR 07-JUN-1995; 95US-0482711.
XX
PR 16-JAN-1995; 95US-0485243.
XX
PR 16-JAN-1998; 98US-0007999.
XX
PR 16-JAN-1998; 98US-0008172.
XX
PR 20-JAN-1998; 98US-0009620.

XX
PA (NICH) NICHOLS S E.

XX
PI Nichols SE;

XX
DR WPI; 2002-414332/44.

PT Glucosyltransferase B or D protein useful for producing a glucan useful
PT as substitutes for and additions to modified starch and latexes in
PT paper manufacture, comprises mutations in specific positions -
XX
PS Claim 36; Page -; 44pp; English.

CC The invention an isolated protein comprising a glucosyltransferase
CC (GTF) B polypeptide having changes at position from 1448V, D457N,
CC D567T, K1014T, D457N/D567T, D457N/D571K, D567T/D571K,
CC D567T/D571K/K1014T, 1448V/D457N/D567T/D571K/K790/K1014T,
CC Y169A/Y170A/Y171A, and K779Q or a GTF D polypeptide having
CC changes at positions from T589D, T589E, N471D, N471D/T589D, and
CC N471D/T589E. Also included are a glucan produced by the GTF mutant,
CC an isolated polynucleotide which encodes P1 or P2, or its complementary
CC polynucleotide, a ribonucleic acid sequence encoding the GTF mutant,

CC an expression cassette comprising the polynucleotide operably linked to a
CC promoter, a vector comprising the expression cassette, host cell
CC introduced with the vector, a transgenic plant comprising the
CC vector, a seed or tuber from the transgenic plant, a paper sizing and/or
CC coating composition comprising a glucan produced in a plant transformed
CC with a gene encoding the mutant GTF, wild type or, starch, a latex,
CC thermoplastic molecule or their combinations and/or vacuole or a maize line
CC deficient in starch biosynthesis, transformed with a gene encoding a
CC glucosyltransferase B or D enzyme, wild-type or mutant and a paper
CC comprising the glucan (paper sizing/coating agent). The vector is useful
CC for producing a glucan in a plant. The method comprises transforming a
CC plant cell with the vector, growing the plant cell under plant growing
CC conditions to produce a regenerated plant and inducing expression of the
CC polynucleotide for a time sufficient to produce the glucan in the
CC regenerated plant, where the vector contains a transit sequence from
CC ribulose biphosphate carboxylase small subunit, waxy, brittle-1 and
CC chlorophyll AB binding protein to produce a transgenic plant, and glucan
CC is produced in the amyloplast of potato or the vacuole of sugar beet.
CC Glucans are useful as substitutes for and additions to modified starch
CC and latexes in paper manufacture. Unlike prior art techniques, which
CC require input materials that produce chemical effluents, paper
CC manufacture utilizing the glucan produced by GTF, which utilizes
CC biologically produced input materials, is more cost-effective and
CC environmentally friendly. Moreover, glucans also exhibit thermoplastic
CC properties and impart gloss to the paper during coating step.
CC The present sequence represents a GTFD mutant of the invention.
CC Note: The present sequence is not shown in the specification but
CC was created by the indexer using the GTFD sequence appearing as AAU98029
CC and the information in claim 36.

XX
SQ Sequence 1430 AA;

Query Match 94.4%; Score 101; DB 23; Length 1430;
Best Local Similarity 95.5%; Pred. No. 5.7e-09;
Matches 21; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Db 1 MANYIFIRAHDSVQVYIAKII 22
574 MANYIFIRAHDSVQVYIAKII 595

RESULT 7
AAR91047
ID AAR91047 standard; Protein; 1577 AA.

XX
AC AAR91047;

XX
DT 22-MAY-1996 (first entry)

XX
DE Alpha-D-glucosyltransferase.

XX
KW Alpha-D-glucosyltransferase; primer-independent; soluble glucan,
KW sucrose; transgenic plant; cloning; Escherichia coli;
KW phage lambda-C13; vector; plasmid pGS301; plasmid pGS302;
KW gene transfer; crop improvement; storage carbohydrate; pasture;
KW feedstuff; senescence; dextran; food; pharmaceutical.

XX
OS Streptococcus salivarius strain ATCC 25975.

XX
PN MO9606173-A1.

XX
PD 29-FEB-1996.

XX
PF 24-AUG-1995; 95WO-AU00527.

XX
PR 24-AUG-1994; 94AU-0007643.

XX
PA (GTFE) GIFFARD P M.
XX (JACQ) JACQUES N A.
XX (SIMP) SIMPSON C L.

PI Giffard PM, Jacques NA, Simpson CL;

XX WPI: 1996-151376/15.
 DR N-PSDB; AATT13139.
 XX
 PT Plants contg. new bacterial DNA encoding glucosyl transferase
 PT activity - retain higher levels of stored carbohydrate(s) in a form
 PT readily digestible by ruminants
 XX
 PS Claim 4; Page 16-20; 31pp; English.
 XX
 CC The sequence represents an alpha-D-glucosyltransferase from
 CC Streptococcus salivarius. The enzyme is primer-independent, and
 CC produces soluble glucan from sucrose. A gene encoding the enzyme
 CC may be cloned and expressed in Escherichia coli using a subclone
 CC of phage lambda-C13, e.g. plasmid pSG501 or plasmid pSG502. The
 CC DNA may also be expressed in a transgenic plant, to improve the
 CC level of stored carbohydrate in a pasture plant which normally
 CC contains low levels, or to prevent degradation of stored carbohydrate
 CC during plant senescence. Dextran may be isolated from the plant, for
 CC use as a food binder or pharmaceutical additive. Primer independence
 CC ensures that the enzyme will be functional in plants. The glucan is
 CC poorly degraded in plants but easily degraded by bacteria in the rumen
 CC of grazing livestock.
 XX
 SQ Sequence 1577 AA;
 Query Match 83.2%; Score 89; DB 17; Length 1577;
 Best Local Similarity 81.0%; Pred. No. 9.4e-07;
 Matches 17; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
 QY 2 ANYIFRADSEYQVYIAKII 22
 DB 660 ANYIFRADSEYQVYIAKII 680
 RESULT 8
 ABB98574
 ID ABB98574 standard; protein; 2835 AA.
 XX
 AC ABB98574;
 XX
 DT 14-JAN-2003 (first entry)
 XX
 DB Dextran saccharase, DSRE.
 XX
 KM Dextran saccharase; enzyme; cytostatic; dermatological; antieborrheic;
 KM DSR-E; glycosyl transferase; dextran; prebiotic; pharmaceutical;
 KM microflora regulation; intestinal transit; mineral assimilation;
 KM colon cancer; acne; dandruff; body odour.
 XX
 OS Leuconostoc mesenteroides NRRL B-1299.
 XX
 FH Key
 FH Peptide Location/Qualifiers
 FT 1..40
 FT /label= Signal_peptide
 FT 41..2835
 FT /label= Mature_protein
 FT 423..439
 FT /label= SEQ_ID_6
 FT /note= "This sequence is specifically claimed in Claim 7"
 FT 478..501
 FT /label= SEQ_ID_7
 FT /note= "This sequence is specifically claimed in Claim 7"
 FT 519..539
 FT /label= SEQ_ID_8
 FT /note= "This sequence is specifically claimed in Claim 7"
 FT 560..571
 FT /label= SEQ_ID_9
 FT /note= "This sequence is specifically claimed in Claim 7"
 FT 631..645
 FT /label= SEQ_ID_10
 FT /note= "This sequence is specifically claimed in Claim 7"
 FT 1014..1021
 FT Peptide

FT /label= SEQ_ID_11
 FT /note= "This sequence is specifically claimed in Claim 7"
 FT 2120..2138
 FT /label= SEQ_ID_12
 FT /note= "This sequence is specifically claimed in Claim 7"
 FT 2161..2184
 FT /label= SEQ_ID_13
 FT /note= "This sequence is specifically claimed in Claim 7"
 FT 2202..2214
 FT /label= SEQ_ID_14
 FT /note= "This sequence is specifically claimed in Claim 7"
 FT 2243..2250
 FT /label= SEQ_ID_15
 FT /note= "This sequence is specifically claimed in Claim 7"
 FT 2315..2322
 FT /label= SEQ_ID_16
 FT /note= "This sequence is specifically claimed in Claim 7"
 FT 2689..2696
 FT /label= SEQ_ID_17
 FT /note= "This sequence is specifically claimed in Claim 7"
 FT 1981..1142
 FT /note= "Catalytic domain, SEQ ID 1. This sequence is
 FT specifically claimed in Claim 4"
 XX
 PN FR2822163-A1.
 XX
 PD 20-SEP-2002.
 XX
 PF 19-DEC-2001; 2001FR-0016495.
 XX
 PR 16-MAR-2001; 2001FR-0003631.
 XX
 PA (CNRS) CNRS CENT NAT RECH SCI.
 XX
 PI Bozonnet SAM, Renaud SMMC, Willemot RML, Monsan PEF.
 XX
 DR WPI: 2002-715213/78.
 DR N-PSDB; ABQ80961, ABQ80962.
 XX
 PT New glycosyl transferase enzymes, containing glucan bonding and
 PT catalytic domains and producing alpha-(1-2) branched dextran, useful
 PT in probiotic, pharmaceutical or cosmetic compositions
 XX
 PS Claim 6; Page 65-74; 82pp; French.
 XX
 CC The present sequence is a novel dextran saccharase, DSR-E, from
 CC Leuconostoc mesenteroides NRRL B-1299. The saccharase has glycosyl
 CC transferase activity suitable for producing dextran having alpha(1-2)
 CC branches from sucrose, p-nitrophenyl-alpha-D-glucopyranoside,
 CC alpha-fluoroglucose, alpha-D-glucopyranoside-alpha-D-sorbidopyranoside or
 CC alpha-D-galactopyranosyl-sucrose. The dextran saccharase is useful in
 CC prebiotic, pharmaceutical or cosmetic compositions. The dextran and
 CC related compounds having alpha(1-2) bonds, produced using DSR-E, may be
 CC involved in signalling/cellular recognition processes in vivo
 CC (specifically in regulation of microflora in the intestines or on the
 CC skin), and are potentially useful for improving intestinal transit,
 CC increasing assimilation of minerals (e.g. calcium and/or magnesium),
 CC preventing cancer of the colon and combating skin problems such as acne,
 CC dandruff and body odour.
 XX
 SQ Sequence 2835 AA;
 Query Match 83.2%; Score 89; DB 23; Length 2835;
 Best Local Similarity 90.0%; Pred. No. 1.9e-06;
 Matches 18; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 QY 3 NYIFRADSEYQVYIAKII 22
 DB 630 NYIFRADSEYQVYIAKII 649
 RESULT 9
 AAN80055

ID		AAU980055standard; Protein; 1527 AA.
XX AC	AAU980055;	
XX XX		
DT DT	30-JUL-2002 (first entry)	
XX DE	Leuconostoc mesenteroides dextranucrase.	
DE XX	Dextranucrase; yoghurt; curd; cheese; fermented milk; infant formulae; pet food; vitamin; oral vaccine; enzyme.	
KM KM	leuconostoc mesenteroides.	
OS OS	EPI2011J31-A1.	
PN PN	02-MAY-2002.	
XX PD	23-OCT-2000; 2000EP-0123012.	
PF PF	23-OCT-2000; 2000BP-0123012.	
XX PR	(NEST) SOC PROD NESTLE SA.	
XX PA	Bauché A., De Maleprade D., Duboc P., Neubauer H., Zink R. WPLI, 2002-373873/41.	
DR DR	N-PBDB; ABK50932.	
PT PT	Preparing fermented food products, e.g. yoghurt, using microorganisms that have dextransucrose bound to the surface, provide attractive and uniform texture -	
XX PS	Disclosure; Page 14-20; 38pp; English.	
XX XX	The invention relates to preparation of a fermented food product that includes treating the microorganism (A) to be used for fermentation with a dextransucrase so that this binds to the cell walls of (A). The method is used to produce yoghurt, curd, cheese or other fermented milk products, ice cream, fermented cereal products, water-based jellies, infant formulae and pet foods. Dextransucrase can also be bound to e.g. vitamins or oral vaccines for delivering these to foods, using (A) as carrier. The method produces foods with attractive and uniform texture, since dextransucrase is evenly distributed and forms a thickening agent, in situ, with the amount formed being controlled by the amount of dextransucrase bound. Milk does not have to be coagulated in order to achieve a good texture, so the amount of acidic by-products formed is reduced. Dextransucrase binds to many different types of cells over wide ranges of temperature and pH and the treated bacteria can be stored in liquid or powdered form. The present sequence represents the amino acid sequence of dextransucrase.	
CC CC	Sequence 1527 AA;	SQ
Query Match:	80.4%; Score 86; DB 23; Length 1527;	
Best Local Similarity	80.0%; Pred. No. 3.le-06;	
Matches 16; Conservative	3; Mismatches 1; Indels 0; Gaps 0;	
QY	3 NYIFIRAHDSVEQTIAKII 22	
	: : :	
Db	654 NYSFVRAHDSVEQTIVAGIV 673	
RESULT 10		
ID ID	AAU98036 standard; Protein; 1475 AA.	
AC AC	AAU98036;	
XX XX		
DT DT	27-AUG-2002 (first entry)	
XX XX	S. mutans glucosyltransferase GTFB mutant D567T/D571K. Glucosyltransferase; GTPB; transgenic plant; paper sizing;	

KW	coating composition; glucan; starch; latex; thermoplastic molecule;
KV	amyloplast; vacuole; paper manufacture; mutant; mutin.
XX	
OS	Streptococcus mutans.
XX	Synthetic.
FH	
Key	Location/Qualifiers
FT	Misc-difference 567
FT	/note= "Wild-type Asp substituted by Thr"
FT	
FT	Misc-difference 571
FT	/note= "Wild-type Asp substituted by Lys"
XX	
US	US2002031826-A1.
PD	
14-MAR-2002.	
PR	
19-DEC-2000; 2000US-0740274.	
XX	
PR	
11-DEC-1998; 98US-0210361.	
PR	
07-JUN-1995; 95US-0478704.	
PR	
07-JUN-1995; 95US-0482711.	
PR	
07-JUN-1995; 95US-0485243.	
PR	
16-JAN-1998; 98US-0007999.	
PR	
16-JAN-1998; 98US-0008172.	
PR	
20-JAN-1998; 98US-0009620.	
XX	
(NICH/) NICHOLS S E.	
PI	
Nichols SE;	
PT	
WP1; 2002-414332/44.	
XX	
Glucosyltransferase B or D protein useful for producing a glucan useful	
as substrates for and additions to modified starch and latexes in	
paper manufacture, comprises mutations in specific positions -	
PS	
Claim 36; Page -; 44pp; English.	
XX	
The invention an isolated protein comprising a glucosyltransferase	
(GTF) B polypeptide having changes at position from I448V, D457N,	
D567T, K1014T, D457N/D567T, D457N/D571X, D567T/D571X,	
D567T/K1014T, I448V/D457N/D567T/D571X/K779Q/K1014T,	
Y169A/Y170M/Y171A, and K779Q or a GTF D polypeptide having	
changes at positions from T589D, T589E, N471D, N471D/T589D, and	
N471D/T589E. Also included are a glucan produced by the GTF mutant,	
an isolated polynucleotide which encodes P1 or P2, or its complementary	
polynucleotide, a ribonucleic acid sequence encoding the GTF mutant,	
an expression cassette comprising the polynucleotide operably linked to a	
promoter, a vector comprising the expression cassette, host cell	
introduced with the vector, a transgenic plant comprising the	
vector, a seed or tuber from the transgenic plant, a paper sling and/or	
coating composition comprising a glucan produced in a plant transformed	
with a gene encoding the mutant GTF, wild type or, starch, a latex,	
thermoplastic molecule or their combinations or glucan and starch where	
the glucan is produced in the amyloplast and/or vacuole or a maize line	
deficient in starch biosynthesis, transformed with a gene encoding a	
glucosyltransferase B or D enzyme, wild-type or mutant and a paper	
comprising the glucan (paper sizing/coating agent). The vector is useful	
for producing a glucan in a plant. The method comprises transforming a	
plant cell with the vector, growing the plant cell under plant growing	
conditions to produce a regenerated plant and inducing expression of the	
polynucleotide for a time sufficient to produce the glucan in the	
regenerated plant, where the vector contains a transit sequence from	
ribulose biphosphate carboxylase small subunit, waxy, brittle-1 and	
chlorophyll AB binding protein to produce a transgenic plant, and glucan	
is produced in the amyloplast of potato or the vacuole of sugar beet.	
Glucans are useful as substitutes for and additions to modified starch	
and latexes in paper manufacture. Unlike prior art techniques, which	
require input materials that produce chemical effluents, paper	
manufacture utilizing the glucan produced by GTF, which utilizes	
biologically produced input materials, is more cost-effective and	
environmentally friendly. Moreover, glucans also exhibit thermoplastic	
properties and impart gloss to the paper during coating step.	

CC The present sequence represents a GTFB mutant of the invention.
CC Note: The present sequence is not shown in the specification but
CC was created by the indexer using the GTFB sequence appearing as AAU98027
CC and the information in claim 36.
XX
SQ Sequence 1475 AA;
Query Match 78.5%; Score 84; DB 23; Length 1475;
Best Local Similarity 85.0%; Pred. No. 6.9e-06;
Matches 17; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
QY 3 NYIFRAHDSVQTVIAKII 22
DB 554 SYSFIRAHDSVQTVIAKII 573
RESULT 11
AAU98037
ID AAU98037 standard; Protein: 1475 AA.
XX
AC AAU98037;
DT 27-AUG-2002 (first entry)
XX
DE S. mutans glucosyltransferase GTFB mutant D567T/D571K/K1014T.
XX
KW Glucosyltransferase; GTFB; transgenic plant; paper sizing;
KW coating composition; glucan; starch; latex; thermoplastic molecule;
KW amyloplast; vacuole; paper manufacture; mutant; muten.
XX
OS Streptococcus mutans.
OS Synthetic.
XX
XX Key Location/Qualifiers
FT Misc-difference 567 /note= "Wild-type Asp substituted by Thr"
FT Misc-difference 571 /note= "Wild-type Asp substituted by Lys"
FT Misc-difference 1014 /note= "Wild-type Lys substituted by Thr"
XX
XX US2002031826-A1.
XX
XX 14-MAR-2002.
XX
XX 19-DEC-2000; 2000US-0740274.
XX
XX 11-DEC-1998; 98US-0210361.
XX 07-JUN-1995; 95US-0478704.
XX 07-JUN-1995; 95US-0482711.
XX 07-JUN-1995; 95US-0485243.
XX 16-JAN-1998; 98US-0007999.
XX 16-JAN-1998; 98US-0008172.
XX 20-JAN-1998; 98US-0009620.
XX
XX (NICH/) NICHOLS S E.
XX
XX Nichols SE;
XX
XX WPI; 2002-414332/44.
XX
XX
XX Glucosyltransferase B or D protein useful for producing a glucan useful
PT as substitutes for and additions to modified starch and latexes in
PT paper manufacture, comprises mutations in specific positions -
XX
XX Claim 36; Page -: 44pp; English.
XX
XX The invention an isolated protein comprising a glucosyltransferase
CC (GTF) B polypeptide having changes at position from I448V, D457N,
CC D567T, K1014T, D457N/D567T, D457N/D571K, D567T/D571K,
CC D567T/D571K/K1014T, I448V/D457N/D567T/D571K/K1014T,
CC Y169K/Y170A/Y171A, and K779Q or a GTF D polypeptide having
CC changes at positions from T589D, T589E, N471D, N471D/T589D, and

CC N471D/T589E. Also included are a glucan produced by the GTF mutant,
CC an isolated polynucleotide which encodes P1 or P2, or its complementary
CC polynucleotide, a ribonucleic acid sequence encoding the GTF mutant,
CC an expression cassette comprising the polynucleotide operably linked to a
CC promoter, a vector comprising the expression cassette, host cell
CC introduced with the vector, a transgenic plant comprising the
CC vector, a seed or tuber from the transgenic plant, a paper sizing and/or
CC coating composition comprising a glucan produced in a plant transformed
CC with a gene encoding the mutant GTF, wild type or, starch, a latex,
CC thermoplastic molecule or their combinations or glucan and starch where
CC the glucan is produced in the amyloplast and/or vacuole or a maize line
CC deficient in starch biosynthesis, transformed with a gene encoding a
CC glucosyltransferase B or D enzyme, wild-type or mutant and a paper
CC comprising the glucan (paper sizing/coating agent). The vector is useful
CC for producing a glucan in a plant. The method comprises transforming a
CC plant cell with the vector, growing the plant cell under plant growing
CC conditions to produce a regenerated plant and inducing expression of the
CC polynucleotide for a time sufficient to produce the glucan in the
CC regenerated plant, where the vector contains a transit sequence from
CC ribulose biphosphate carboxylase small subunit, waxy, brittle-1 and
CC chlorophyll AB binding protein to produce a transgenic plant, and glucan
CC is produced in the amyloplast of potato or the vacuole of sugar beet.
CC Glucans are useful as substitutes for and additions to modified starch
CC and latexes in paper manufacture. Unlike prior art techniques, which
CC require input materials that produce chemical effluents, paper
CC manufacture utilizing the glucan produced by GTF, which utilizes
CC biologically produced input materials, is more cost-effective and
CC environmentally friendly. Moreover, glucans also exhibit thermoplastic
CC properties and impart gloss to the paper during coating step.
CC The present sequence represents a GTFB mutant of the invention.
CC Note: The present sequence is not shown in the specification but
CC was created by the indexer using the GTFB sequence appearing as AAU98027
CC and the information in claim 36.
XX
SQ Sequence 1475 AA;
Query Match 78.5%; Score 84; DB 23; Length 1475;
Best Local Similarity 85.0%; Pred. No. 6.9e-06;
Matches 17; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
QY 3 NYIFRAHDSVQTVIAKII 22
DB 554 SYSFIRAHDSVQTVIAKII 573
RESULT 12
AAU98038
ID AAU98038 standard; Protein: 1475 AA.
XX
XX AAU98038;
XX
XX 27-AUG-2002 (first entry)
XX
XX S. mutans GTFB mutant I448V/D457N/D567T/D571K/K779Q/K1014T.
XX
XX
XX Glucosyltransferase; GTFB; transgenic plant; paper sizing;
KW coating composition; glucan; starch; latex; thermoplastic molecule;
KW amyloplast; vacuole; paper manufacture; mutant; muten.
XX
XX Streptococcus mutans.
OS Synthetic.
OS
XX Key Location/Qualifiers
FT Misc-difference 448 /note= "Wild-type Ile substituted by Val"
FT Misc-difference 457 /note= "Wild-type Asp substituted by Asn"
FT Misc-difference 567 /note= "Wild-type Asp substituted by Thr"
FT Misc-difference 571 /note= "Wild-type Asp substituted by Lys"
FT Misc-difference 779 /note= "Wild-type Asp substituted by Lys"
FT Misc-difference 779 /note= "Wild-type Lys substituted by Gln"

FT Misc-difference 1014 /note= "Wild-type Lys substituted by Thr"
 XX
 XX US2002031826-A1.
 XX
 XX 14-MAR-2002.
 XX
 XX 19-DEC-2000; 2000US-0740274.
 XX
 XX 11-DEC-1998; 98US-0210361.
 XX 07-JUN-1995; 95US-0478704.
 XX 07-JUN-1995; 95US-0482711.
 XX 07-JUN-1995; 95US-0485243.
 XX 16-JAN-1998; 98US-0007999.
 XX 16-JAN-1998; 98US-0008172.
 XX 20-JAN-1998; 98US-0009620.
 XX
 XX (NICH/) NICHOLS S E.
 XX
 XX Nichols SE;
 XX
 XX WPI; 2002-414332/44.
 XX
 XX Glucosyltransferase B or D protein useful for producing a glucan useful
 XX as substitutes for and additions to modified starch and latexes in
 XX paper manufacture, comprises mutations in specific positions -
 XX
 XX Claim 36; Page -; 44pp; English.
 XX
 XX The invention an isolated protein comprising a glucosyltransferase
 XX (GTF) B polypeptide having changes at position from 1448V, D457N,
 XX D567T, K1014T, D457N/D567T, D457N/D571K, D567T/D571K,
 XX D567T/D571K/K1014T, I448V/D457N/D567T/D571K/K779Q/K1014T,
 XX Y169A/Y170A/Y171A, and K779Q or a GTF D polypeptide having
 XX changes at positions from T589D, T589E, N471D, N471D/T589D, and
 XX N471D/T589E. Also included are a glucan produced by the GTF mutant,
 XX an isolated polynucleotide which encodes P1 or P2, or its complementary
 XX polynucleotide, a ribonucleic acid sequence encoding the GTF mutant,
 XX an expression cassette comprising the polynucleotide operably linked to a
 XX promoter, a vector comprising the expression cassette, host cell
 XX introduced with the vector, a transgenic plant comprising the
 XX vector, a seed or tuber from the transgenic plant, a paper sizing and/or
 XX coating composition comprising a glucan produced in a plant transformed
 XX with a gene encoding the mutant GTF, wild type or, starch, a latex,
 XX thermoplastic molecule or their combinations or glucan and starch where
 XX the glucan is produced in the amyloplast and/or vacuole or a maize line
 XX deficient in starch biosynthesis, transformed with a gene encoding a
 XX glucosyltransferase B or D enzyme, wild-type or mutant and a paper
 XX comprising the glucan (paper sizing/coating agent). The vector is useful
 XX for producing a glucan in a plant. The method comprises transforming a
 XX plant cell with the vector, growing the plant cell under plant growing
 XX conditions to produce a regenerated plant and inducing expression of the
 XX polynucleotide for a time sufficient to produce the glucan in the
 XX regenerated plant, where the vector contains a transit sequence from
 XX ribulose biphosphate carboxylase small subunit, waxy, brittle-1 and
 XX chlorophyll AB binding protein to produce a transgenic plant, and glucan
 XX is produced in the amyloplast of potato or the vacuole of sugar beet.
 XX Glucans are useful as substitutes for and additions to modified starch
 XX and latexes in paper manufacture. Unlike prior art techniques, which
 XX require input materials that produce chemical effluents, paper
 XX manufacture utilizing the glucan produced by GTF, which utilizes
 XX biologically produced input materials, is more cost-effective and
 XX environmentally friendly. Moreover, glucans also exhibit thermoplastic
 XX properties and impart gloss to the paper during coating step.
 XX The present sequence represents a GTFB mutant of the invention.
 XX Note: The present sequence is not shown in the specification but
 XX and the information in claim 36.
 XX
 XX Sequence 1475 AA;
 XX
 XX Query Match 78.5%; Score 84; DB 23; Length 1475;
 XX Best Local Similarity 85.0%; Pred. No. 6.9e-06;

Matches 17; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
 QY 3 NYIFRAHSEVQTVIAKII 22
 : |||||
 Db 554 STSFIRAHSEVQTVIAKII 573
 RESULT 13
 AAU98032
 ID AAU98032 standard; Protein, 1475 AA.
 XX
 XX AAU98032;
 XX
 XX 27-AUG-2002 (first entry)
 XX
 XX S. mutans glucosyltransferase GTFB mutant D567T.
 XX
 XX Glucosyltransferase; GTFB; transgenic plant; paper sizing;
 XX coating composition; glucan; starch; latex; thermoplastic molecule;
 XX amyloplast; vacuole; paper manufacture; mutant; muten.
 XX
 XX Streptococcus mutans.
 XX Synthetic.
 XX
 XX Key Location/Qualifiers
 XX
 XX Misc-difference 567 /note= "Wild-type Asp substituted by Thr"
 XX
 XX US2002031826-A1.
 XX
 XX 14-MAR-2002.
 XX
 XX 19-DEC-2000; 2000US-0740274.
 XX
 XX 11-DEC-1998; 98US-0210361.
 XX 07-JUN-1995; 95US-0478704.
 XX 07-JUN-1995; 95US-0482711.
 XX 07-JUN-1995; 95US-0485243.
 XX 16-JAN-1998; 98US-0007999.
 XX 16-JAN-1998; 98US-0008172.
 XX 20-JAN-1998; 98US-0009620.
 XX
 XX (NICH/) NICHOLS S E.
 XX
 XX Nichols SE;
 XX
 XX WPI; 2002-414332/44.
 XX
 XX Glucosyltransferase B or D protein useful for producing a glucan useful
 XX as substitutes for and additions to modified starch and latexes in
 XX paper manufacture, comprises mutations in specific positions -
 XX
 XX Claim 36; Page -; 44pp; English.
 XX
 XX The invention an isolated protein comprising a glucosyltransferase
 XX (GTF) B polypeptide having changes at position from 1448V, D457N,
 XX D567T, K1014T, D457N/D567T, D457N/D571K, D567T/D571K,
 XX D567T/D571K/K1014T, I448V/D457N/D567T/D571K/K779Q/K1014T,
 XX Y169A/Y170A/Y171A, and K779Q or a GTF D polypeptide having
 XX changes at positions from T589D, T589E, N471D, N471D/T589D, and
 XX N471D/T589E. Also included are a glucan produced by the GTF mutant,
 XX an isolated polynucleotide which encodes P1 or P2, or its complementary
 XX polynucleotide, a ribonucleic acid sequence encoding the GTF mutant,
 XX an expression cassette comprising the polynucleotide operably linked to a
 XX promoter, a vector comprising the expression cassette, host cell
 XX introduced with the vector, a transgenic plant comprising the
 XX vector, a seed or tuber from the transgenic plant, a paper sizing and/or
 XX coating composition comprising a glucan produced in a plant transformed
 XX with a gene encoding the mutant GTF, wild type or, starch, a latex,
 XX thermoplastic molecule or their combinations or glucan and starch where
 XX the glucan is produced in the amyloplast and/or vacuole or a maize line
 XX deficient in starch biosynthesis, transformed with a gene encoding a
 XX glucosyltransferase B or D enzyme, wild-type or mutant and a paper

comprising the glucan (paper sizing/coating agent). The vector is useful for producing a glucan in a plant. The method comprises transforming a plant cell with the vector, growing the plant cell under plant growing conditions to produce a regenerated plant and inducing expression of the polynucleotide for a time sufficient to produce the glucan in the regenerated plant, where the vector contains a transit sequence from chlorophyll AB binding protein to produce a transgenic plant, and glucan is produced in the amyloplast of potato or the vacuole of sugar beet. Glucans are useful as substitutes for and additions to modified starch and latexes in paper manufacture. Unlike prior art techniques, which require input materials that produce chemical effluents, paper manufacture utilizing the glucan produced by GTF, which utilizes biologically produced input materials, is more cost-effective and environmentally friendly. Moreover, glucans also exhibit thermoplastic properties and impart gloss to the paper during coating step. The present sequence represents a GTFB mutant of the invention. Note: The present sequence is not shown in the specification but was created by the indexer using the GTFB sequence appearing as AAU98027 and the information in claim 36.

Sequence 1475 AA;
Query Match 72.9%; Score 78; DB 23; Length 1475;
Best Local Similarity 80.0%; Pred. No. 8.4e-05;
Matches 16; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
3 NYIFIRADSEVQVIADII 22
554 SYSFIRADSEVQVIADII 573

RESULT 14
AAU98034
ID AAU98034 standard; Protein; 1475 AA.

AAU98034;
27-AUG-2002 (first entry)
S. mutans glucosyltransferase GTFB mutant D457N/D567T.
Glucosyltransferase; GTFB; transgenic plant; paper sizing;
coating composition; glucan; starch; latex; thermoplastic molecule;
amyloplast; vacuole; paper manufacture; mutant; mutlein.
Streptococcus mutans.
Synthetic.

Key Location/Qualifiers
Misc-difference 457 /note= "Wild-type Asp substituted by Asn"
Misc-difference 567 /note= "Wild-type Asp substituted by Thr"

US2002031826-A1.
14-MAR-2002.
19-DEC-2000; 2000US-0740274.
11-DEC-1998; 98US-0210361.
07-JUN-1995; 95US-0478704.
07-JUN-1995; 95US-0482711.
07-JUN-1995; 95US-0485243.
16-JAN-1998; 98US-0007999.
16-JAN-1998; 98US-0008172.
20-JAN-1998; 98US-0009620.

(NICH/) NICHOLS S E.
Nichols SE;
XX

WPI; 2002-414332/44.
Glucosyltransferase B or D protein useful for producing a glucan useful as substitutes for and additions to modified starch and latexes in paper manufacture, comprises mutations in specific positions -
Claim 36; Page -; 44pp; English.

The invention an isolated protein comprising a glucosyltransferase (GTF) B polypeptide having changes at position from I448V, D457N, D567T, K1014T, D457N/D567T, D457N/D571K, D567T/D571K, D567T/D571K/K1014T, I448V/D457N/D567T/D571K/K799Q/K1014T, Y169A/Y170A/Y171A, and K779Q or a GTF D polypeptide having changes at positions from T589D, T589E, N471D, N471D/T589D, and N471D/T589E. Also included are a glucan produced by the GTF mutant, an isolated polynucleotide which encodes P1 or P2, or its complementary polynucleotide, a ribonucleic acid sequence encoding the GTF mutant, an expression cassette comprising the polynucleotide operably linked to a promoter, a vector comprising the expression cassette, host cell introduced with the vector, a transgenic plant comprising the vector, a seed or tuber from the transgenic plant, a plant transformed with a gene encoding the mutant GTF, wild type or, starch, a latex, thermoplastic molecule or their combinations or glucan and starch where the glucan is produced in the amyloplast and/or vacuole or a maize line deficient in starch biosynthesis, transformed with a gene encoding a glucosyltransferase B or D enzyme, wild-type or mutant and a paper comprising the glucan (paper sizing/coating agent). The vector is useful for producing a glucan in a plant. The method comprises transforming a plant cell with the vector, growing the plant cell under plant growing conditions to produce a regenerated plant and inducing expression of the polynucleotide for a time sufficient to produce the glucan in the regenerated plant, where the vector contains a transit sequence from chlorophyll AB binding protein to produce a transgenic plant, and glucan is produced in the amyloplast of potato or the vacuole of sugar beet. Glucans are useful as substitutes for and additions to modified starch and latexes in paper manufacture. Unlike prior art techniques, which require input materials that produce chemical effluents, paper manufacture utilizing the glucan produced by GTF, which utilizes biologically produced input materials, is more cost-effective and environmentally friendly. Moreover, glucans also exhibit thermoplastic properties and impart gloss to the paper during coating step. The present sequence represents a GTFB mutant of the invention. Note: The present sequence is not shown in the specification but was created by the indexer using the GTFB sequence appearing as AAU98027 and the information in claim 36.

Sequence 1475 AA;
Query Match 72.9%; Score 78; DB 23; Length 1475;
Best Local Similarity 80.0%; Pred. No. 8.4e-05;
Matches 16; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
3 NYIFIRADSEVQVIADII 22
554 SYSFIRADSEVQVIADII 573

RESULT 15
AAU98035
ID AAU98035 standard; Protein; 1475 AA.

AAU98035;
27-AUG-2002 (first entry)

S. mutans glucosyltransferase GTFB mutant D457N/D571K.
Glucosyltransferase; GTFB; transgenic plant; paper sizing;
coating composition; glucan; starch; latex; thermoplastic molecule;
amyloplast; vacuole; paper manufacture; mutant; mutlein.
XX

OS Streptococcus mutans.
 OS Synthetic.
 XX
 XX
 PH Key Location/Qualifiers
 FT Misc-difference 457
 FT /note= "Wild-type Asp substituted by Asn"
 FT Misc-difference 571
 FT /note= "Wild-type Asp substituted by Lys"
 XX
 XX US2002031826-A1.
 XX
 XX
 PD 14-MAR-2002.
 XX
 XX
 PP 19-DEC-2000; 2000US-0740274.
 XX
 XX 11-DEC-1998; 98US-0210361.
 PR 07-JUN-1995; 95US-0478704.
 PR 07-JUN-1995; 95US-0482711.
 PR 07-JUN-1995; 95US-0485243.
 PR 16-JAN-1998; 98US-0007999.
 PR 16-JAN-1998; 98US-0008172.
 PR 20-JAN-1998; 98US-0009620.
 XX
 XX (NICH/) NICHOLS S E.
 XX
 XX Nichols SE;
 XX
 DR WPI; 2002-414332/44.
 XX
 XX
 PT Glucosyltransferase B or D protein useful for producing a glucan useful
 PT as substitutes for and additions to modified starch and latexes in
 PT paper manufacture, comprises mutations in specific positions -
 XX
 XX
 PS Claim 36; Page -; 44pp; English.
 XX
 XX
 CC The invention an isolated protein comprising a glucosyltransferase
 CC (GTF) B polypeptide having changes at position from I448V, D457N,
 CC D567T, K1014T, D457N/D567T, D457N/D571K, D567T/D571K,
 CC D567T/D571K/K1014T, I448V/D457N/D567T/D571K/K779Q/K1014T,
 CC Y169A/Y170A/Y171A, and K779Q or a GTF D polypeptide having
 CC changes at positions from T589D, T589E, N471D, N471D/T589D, and
 CC N471D/T589E. Also included are a glucan produced by the GTF mutant,
 CC an isolated polynucleotide which encodes P1 or P2, or its complementary
 CC polynucleotide, a ribonucleic acid sequence encoding the GTF mutant,
 CC an expression cassette comprising the polynucleotide operably linked to a
 CC promoter, a vector comprising the expression cassette, host cell
 CC introduced with the vector, a transgenic plant comprising the
 CC vector, a seed or tuber from the transgenic plant, a paper sizing and/or
 CC coating composition comprising a glucan produced in a plant transformed
 CC with a gene encoding the mutant GTF, wild type or, starch, a latex,
 CC thermoplastic molecule or their combinations or glucan and starch where
 CC the glucan is produced in the amyloplast and/or vacuole or a maize line
 CC deficient in starch biosynthesis, transformed with a gene encoding a
 CC glucosyltransferase B or D enzyme, wild-type or mutant and a paper
 CC comprising the glucan (paper sizing/coating agent). The vector is useful
 CC for producing a glucan in a plant. The method comprises transforming a
 CC plant cell with the vector, growing the plant cell under plant growing
 CC conditions to produce a regenerated plant and inducing expression of the
 CC polynucleotide for a time sufficient to produce the glucan in the
 CC regenerated plant, where the vector contains a transit sequence from
 CC ribulose biphosphate carboxylase small subunit, waxy, brittle-1 and
 CC chlorophyll AB binding protein to produce a transgenic plant, and glucan
 CC is produced in the amyloplast of potato or the vacuole of sugar beet.
 CC Glucans are useful as substitutes for and additions to modified starch
 CC and latexes in paper manufacture. Unlike prior art techniques, which
 CC require input materials that produce chemical effluents, paper
 CC manufacture utilizing the glucan produced by GTF, which utilizes
 CC biologically produced input materials, is more cost-effective and
 CC environmentally friendly. Moreover, glucans also exhibit thermoplastic
 CC properties and impart gloss to the paper during coating step.
 CC The present sequence represents a GTFB mutant of the invention.
 CC Note: The present sequence is not shown in the specification but
 CC was created by the indexer using the GTFB sequence appearing as AAU98027

CC and the information in claim 36.
 XX
 XX
 SQ Sequence 1475 AA;
 Query Match 72.9%; Score 78; DB 23; Length 1475;
 Best Local Similarity 80.0%; Pred. No. 8.4e-05;
 Matches 16; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
 QY 3 NYIFIRAHDSVQTVIAKII 22
 : ||||| : |||||
 Db 554 SYSFIRAHDSVQDLIAKII 573

Search completed: November 13, 2003, 09:38:29
 Job time : 41.3507 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 13, 2003, 09:31:40 ; Search time 13.6588 Seconds
(without alignments)
154.898 Million cell updates/sec

Title: US-09-290-049A-16

Perfect score: 107

Sequence: 1 MANYIFIRAHDSVQTVIAKII 22

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : PIR 76:*

1: PIR1: *
2: PIR2: *
3: PIR3: *
4: PIR4: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	107	100.0	1431	2 A45866	dextranucrase (EC
2	90	84.1	1365	2 A41483	glucosyltransferas
3	89	83.2	1577	2 T30858	glucosyltransferas
4	88	82.2	1449	2 T30857	glucosyltransferas
5	88	82.2	1449	2 T30852	glucosyltransferas
6	87	81.3	1508	2 T31098	probable dextranu
7	87	81.3	1599	2 S22737	glucosyltransferas
8	85	79.4	1290	2 JCS473	dextranucrase (EC
9	78	72.9	1518	2 A44811	glucosyltransferas
10	72	67.3	1475	2 B31335	gtfB protein precu
11	68	63.6	1375	2 UT0345	dextranucrase (EC
12	64	59.8	1592	2 A38175	glucosyltransferas
13	45	42.1	267	2 B84213	hypothetical prote
14	44.5	41.6	593	2 S51946	pyruvate kinase (E
15	44	41.1	1332	2 D82080	mutator Mult prote
16	44	41.1	431	2 S49821	PR22 protein - Ara
17	44	41.1	583	2 T04531	nine-cis-epoxycaro
18	44	40.2	1213	2 T41378	probable helicase
19	44	40.2	288	2 T26383	hypothetical prote
20	43	40.2	933	2 G70166	phosphate zinc prot
21	42	39.3	236	2 E82593	nitric-oxide regulon
22	42	39.3	1153	2 T31080	nitric-oxide synth
23	41.5	38.8	354	2 A70255	immunogenic protei
24	41.5	38.8	493	2 T10054	pyruvate kinase (E
25	41.5	38.8	583	2 T10051	pyruvate kinase (E
26	41	38.3	76	2 JQ1633	HCFP3 protein - hu
27	41	38.3	140	2 F70515	probable lipoprote
28	41	38.3	283	2 AF1276	hypothetical prote
29	41	38.3	368	2 T40115	uv excision repair

30	41	38.3	411	2 A34526	ORF1 protein - Ory
31	41	38.3	411	2 T10373	hypothetical prote
32	41	38.3	602	2 A71256	aspartate-tRNA lig
33	41	38.3	725	2 T17732	helicase-like prot
34	41	38.3	768	2 S52684	probable membrane
35	41	38.3	1165	2 D72496	hypothetical prote
36	41	38.3	1302	2 AC2017	cobalamin biosynth
37	41	38.3	1771	2 S76851	hypothetical prote
38	40.5	37.9	217	1 H69797	rihamogalacturonan
39	40	37.4	344	1 GRYC87	sulfate transport
40	40	37.4	354	2 S65687	(A+T)-stretch-bind
41	40	37.4	382	2 G90540	hypothetical prote
42	40	37.4	417	2 S25627	glucose-1-phosphat
43	40	37.4	554	2 C69161	sensory transducti
44	40	37.4	557	2 S73434	aspartate-tRNA lig
45	40	37.4	586	2 B84271	glutamyl-tRNA synt

ALIGNMENTS

RESULT 1

A45866 dextranucrase (EC 2.4.1.5) precursor - Streptococcus mutans

C/Species: Streptococcus mutans

C/Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 07-May-1999

C/Accession: A45866

R/Honda, O.; Kato, C.; Kuramitsu, H.K.

J. Gen. Microbiol. 136, 2099-2105, 1990

A/Title: Nucleotide sequence of the Streptococcus mutans gtfD gene encoding the glucosyltr

A/Reference number: A45866; MUID:91100958; PMID:2148600

A/Accession: A45866

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-1431 <HON>

A/Cross-references: GB:M29296

C/Superfamily: cpl repeat homology

C/Keywords: glycosyltransferase; hexosyltransferase

F/181-201/Domain: cpl repeat homology <CP1>

F/1127-1146/Domain: cpl repeat homology <CP2>

F/1192-1211/Domain: cpl repeat homology <CP3>

F/1257-1276/Domain: cpl repeat homology <CP4>

F/1277-1297/Domain: cpl repeat homology <CP5>

F/1321-1340/Domain: cpl repeat homology <CP6>

F/1341-1361/Domain: cpl repeat homology <CP6>

F/1385-1404/Domain: cpl repeat homology <CP7>

Query Match 100.0%; Score 107; DB 2; Length 1431;

Best Local Similarity 100.0%; Pred. No. 2.5e-09;

Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MANYIFIRAHDSVQTVIAKII 22

Db 574 MANYIFIRAHDSVQTVIAKII 595

RESULT 2

A41483 glucosyltransferase (EC 2.4.1.-) gtfS precursor - Streptococcus sobrinus

C/Species: Streptococcus sobrinus

C/Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 15-Oct-1999

C/Accession: A41483

R/Gilmore, K.S.; Russell, R.R.B.; Ferretti, J.J.

Infect. Immun. 58, 2452-2458, 1990

A/Title: Analysis of the Streptococcus downei gtfS gene, which specifies a glucosyltransferase

A/Reference number: A41483; MUID:90316665; PMID:2142479

A/Accession: A41483

A/Molecule type: DNA

A/Residues: 1-1365 <GIL>

A/Cross-references: GB:M30943; MUID:9153652; PIDN:AAA26898.1; PID:9153653

C/Genetics: gtfS

A:Gene: gtfS

C/Superfamily: cpl repeat homology

C:Keywords: glycosyltransferase; hexosyltransferase

Query Match 84.1%; Score 90; DB 2; Length 1365;
Best Local Similarity 90.0%; Pred. No. 1.6e-06;
Matches 18; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 NYIFIRAHDSVQTVIAKII 22
DB 539 NYVIFRAHDSVQTVIAKII 558

RESULT 3

T30858
glucosyltransferase - Streptococcus salivarius

C:Species: Streptococcus salivarius

C:Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 22-Oct-1999

C:Accession: T30858

R:Simpson, C.L.; Giffard, P.M.; Jacques, N.A.

Infect. Immun. 63, 609-621, 1995

A:Title: Streptococcus salivarius ATCC 25975 possesses at least two genes coding for pr

A:Reference number: Z20909; MUID:95122197; PMID:7822030

A:Accession: T30858

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-1577 <SIM>

A:Cross-references: EMBL:L35928; NID:9662380; PID:9662381; PIDN:AA041413.1

C:Genetics:

A:Gene: gtfm

Query Match 83.2%; Score 89; DB 2; Length 1577;
Best Local Similarity 81.0%; Pred. No. 2.8e-06;
Matches 17; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 ANYIFRAHDSVQTVIAKII 22
DB 660 ANYIFRAHDSVQTVIANII 680

RESULT 4

T30857
glucosyltransferase - Streptococcus salivarius

C:Species: Streptococcus salivarius

C:Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 22-Oct-1999

C:Accession: T30857

R:Simpson, C.L.; Giffard, P.M.; Jacques, N.A.

Infect. Immun. 63, 609-621, 1995

A:Title: Streptococcus salivarius ATCC 25975 possesses at least two genes coding for pr

A:Reference number: Z20909; MUID:95122197; PMID:7822030

A:Accession: T30857

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-1449 <SIM>

A:Cross-references: EMBL:L35495; NID:9662378; PID:9662379; PIDN:AA041412.1

C:Genetics:

A:Gene: gtfL

Query Match 82.2%; Score 88; DB 2; Length 1449;
Best Local Similarity 72.7%; Pred. No. 3.7e-06;
Matches 16; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 1 MANYIFRAHDSVQTVIAKII 22
DB 607 MANYIFRAHDSVQTVIAKII 628

RESULT 5

T30552
glucosyltransferase N - Streptococcus salivarius (fragment)

C:Species: Streptococcus salivarius

C:Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 22-Oct-1999

C:Accession: T30552

R:Jaaffe, R.I.

submitted to the EMBL Data Library, February 1998

A:Description: Streptococcus salivarius V1477 gtfN.

A:Reference number: Z20854

A:Accession: T30552

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-1449 <JAF>

A:Cross-references: EMBL:AF049609; NID:92935545; PID:92935546; PIDN:AA005156.1

C:Genetics:

A:Gene: gtfN

Query Match 82.2%; Score 88; DB 2; Length 1449;
Best Local Similarity 72.7%; Pred. No. 3.7e-06;
Matches 16; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 1 MANYIFRAHDSVQTVIAKII 22
DB 607 MANYIFRAHDSVQTVIAKII 628

RESULT 6

T31098
probable dextranucrase (EC 2.4.1.5), extracellular - Leuconostoc mesenteroides

C:Species: Leuconostoc mesenteroides

C:Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 11-May-2000

C:Accession: T31098

R:Monchois, V.; Renaud-Simeon, M.; Monsan, P.; Willemot, R.M.

EMS Microbiol. Lett. 159, 307-315, 1998

A:Title: Cloning and sequencing of a gene coding for an extracellular dextranucrase (DS

A:Reference number: Z20981; MUID:98164374; PMID:9503626

A:Accession: T31098

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-1508 <MON>

A:Cross-references: EMBL:AF030129; NID:92766611; PID:92766612; PIDN:AAB95453.1

A:Experimental source: strain NRRL B-1299

C:Genetics:

A:Gene: dsrB

C:Function:

A:Description: produces dextran composed only of alpha(1-6) glucosidic bonds

C:Keywords: glycosyltransferase; hexosyltransferase

Query Match 81.3%; Score 87; DB 2; Length 1508;
Best Local Similarity 85.0%; Pred. No. 5.7e-06;
Matches 17; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 NYIFRAHDSVQTVIAKII 22
DB 636 NYVIFRAHDSVQTVIAKII 655

RESULT 7

S22737
glucosyltransferase (EC 2.4.1.-) S - Streptococcus salivarius

C:Species: Streptococcus salivarius

C:Date: 16-Sep-1992 #sequence_revision 16-Sep-1992 #text_change 21-Jan-2000

C:Accession: S22737; S28810; B44811; S22727

R:Jacques, N. the EMBL Data Library, March 1992

submitted to the EMBL Data Library, March 1992

A:Reference number: S22726

A:Accession: S22737

A:Molecule type: DNA

A:Residues: 1-1599 <JNC>

A:Cross-references: EMBL:Z11872; NID:947530; PIDN:CAA77898.1; PID:947531

A:Experimental source: ATCC 25975

R:Giffard, P.M.; Simpson, C.L.; Milward, C.P.; Jacques, N.A.

J. Gen. Microbiol. 137, 2577-2593, 1991

A:Title: Molecular characterization of a cluster of at least two glucosyltransferase gen

A:Reference number: A44811; MUID:92148377; PMID:1838391

A:Accession: S28810

A:Molecule type: DNA

A:Residues: 1-51 <GIF>

A:Cross-references: EMBL:Z11873

C:Genetics:

A:Gene: gtfK
 C:Superfamily: cpl repeat homology
 C:Keywords: glycosyltransferase; hexosyltransferase
 F:1456-1475/Domain: cpl repeat homology <CPR>

Query Match 81.3%; Score 87; DB 2; Length 1599;
 Best Local Similarity 81.0%; Pred. No. 6.1e-06;
 Matches 17; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 2 ANYFIRAHDSVQTVIAKII 22
 ||:|||||:|||||:|||||:
 DB 573 ATYFIRAHDSVQTVIADII 593

RESULT 8
 JCS473
 dextranase (EC 2.4.1.5) - Leuconostoc mesenteroides
 C:Species: Leuconostoc mesenteroides
 C>Date: 07-Jul-1997 #sequence_revision 29-Aug-1997 #text_change 29-Aug-1997
 C:Accession: JCS473
 R:Monchois, V.; Willemot, R.M.; Renaud-Simeon, M.; Croux, C.; Monsan, P.
 Gene 182, 23-32, 1996
 A:Title: Cloning and sequencing of a gene coding for a novel dextranase from *Leuconostoc*
 A:Reference number: JCS473; MUID:97136886; PMID:8982063
 A:Accession: JCS473
 A:Status: nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-1290 <MON>
 A:Cross-references: GB:U8181
 C:Comment: This enzyme catalyzes the transfer of D-glucopyranosyl units from sucrose onto
 A:Genetics:
 A:Gene: dsrA
 C:Keywords: glycosyltransferase; hexosyltransferase
 F:78-870/Domain: catalytic #status predicted <CAT>
 F:922-1290/Domain: glucan-binding #status predicted <GCB>

Query Match 79.4%; Score 85; DB 2; Length 1290;
 Best Local Similarity 85.0%; Pred. No. 1e-05;
 Matches 17; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 3 NYFIRAHDSVQTVIAKII 22
 ||:|||||:|||||:|||||:
 DB 389 NYFIRAHDSVQTVIADII 408

RESULT 9
 A44811
 glucosyltransferase (EC 2.4.1.-) I - Streptococcus salivarius
 C:Species: Streptococcus salivarius
 C>Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 15-Oct-1999
 C:Accession: A44811; S22726; S28809
 R:Giffard, P.M.; Simpson, C.L.; Milward, C.P.; Jacques, N.A.
 J. Gen. Microbiol. 137, 2577-2593, 1991
 A:Title: Molecular characterization of a cluster of at least two glucosyltransferase genes
 A:Reference number: A44811; MUID:92148377; PMID:1838391
 A:Accession: A44811
 A:Molecule type: DNA
 A:Residues: 1-1518 <GIF>
 A:Cross-references: EMBL:Z11873; NID:947526; PIDN:CAA77900.1; PID:947527
 A:Note: sequence extracted from NCBI backbone (NCBIN:01050, NCBIIP:01052)
 C:Genetics:
 A:Gene: gtfU
 C:Superfamily: cpl repeat homology
 C:Keywords: glycosyltransferase; hexosyltransferase
 F:1307-1326/Domain: cpl repeat homology <CP4>

Query Match 72.9%; Score 78; DB 2; Length 1518;
 Best Local Similarity 70.0%; Pred. No. 0.00018;
 Matches 14; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

OY 3 NYFIRAHDSVQTVIAKII 22
 ||:|||||:|||||:|||||:
 DB 604 NYFIRAHDSVQTVIADII 623

RESULT 10
 B33135
 gtfB protein precursor - Streptococcus mutans

C:Species: Streptococcus mutans
 C>Date: 23-Oct-1990 #sequence_revision 23-Oct-1990 #text_change 15-Oct-1999
 C:Accession: B33135; A33128
 R:Shiroza, T.; Ueda, S.; Kuramitsu, H.K.
 J. Bacteriol. 169, 4263-4270, 1987
 A:Title: Sequence analysis of the gtfB gene from Streptococcus mutans.
 A:Reference number: A33135; MUID:87308013; PMID:3040685
 A:Accession: B33135
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-1475 <SHI>
 A:Cross-references: GB:M17361; NID:9153639; PIDN:AAA88588.1; PID:9153640
 R:Shiroza, T.; Ueda, S.; Kuramitsu, H.K.
 submitted to the Protein Sequence Database, September 1990
 A:Reference number: A33128
 A:Accession: A33128
 A:Status: preliminary; not compared with conceptual translation
 A:Molecule type: DNA
 A:Residues: 1-171,173-641, 'N', 643-1475 <SH2>
 A:Experimental source: strain GS-5
 C:Superfamily: cpl repeat homology
 F:1096-1115/Domain: cpl repeat homology <CP1>
 F:1224-1243/Domain: cpl repeat homology <CP2>
 F:1289-1308/Domain: cpl repeat homology <CP3>
 F:1354-1373/Domain: cpl repeat homology <CP4>
 F:1419-1438/Domain: cpl repeat homology <CP5>

Query Match 67.3%; Score 72; DB 2; Length 1475;
 Best Local Similarity 75.0%; Pred. No. 0.0018;
 Matches 15; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 3 NYFIRAHDSVQTVIAKII 22
 ||:|||||:|||||:|||||:
 DB 554 SYFIRAHDSVQTVIADII 573

RESULT 11
 UT0345
 dextranase (EC 2.4.1.5) precursor - Streptococcus mutans (strain GS-5)
 N:Alternate names: sucrose 6-glucosyltransferase
 C:Species: Streptococcus mutans
 C>Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 24-Sep-1999
 C:Accession: UT0345; C33135
 R:Ueda, S.; Shiroza, T.; Kuramitsu, H.K.
 Gene 69, 101-109, 1988
 A:Title: Sequence analysis of the gtfC gene from Streptococcus mutans GS-5.
 A:Reference number: UT0345; MUID:89137980; PMID:2976010
 A:Accession: UT0345
 A:Molecule type: DNA
 A:Residues: 1-1375 <UED>
 A:Experimental source: GS-5
 R:Shiroza, T.; Ueda, S.; Kuramitsu, H.K.
 J. Bacteriol. 169, 4263-4270, 1987
 A:Title: Sequence analysis of the gtfB gene from Streptococcus mutans.
 A:Reference number: A33135; MUID:87308013; PMID:3040685
 A:Accession: C33135
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-349 <SHI>
 A:Cross-references: GB:M17361
 C:Genetics:
 A:Gene: gtfC
 C:Function:
 A:Description: catalyzes the synthesis of both water-soluble and water-insoluble glucans
 C:Superfamily: cpl repeat homology
 C:Keywords: duplication; glycosyltransferase; hexosyltransferase
 F:1-34/Domain: signal sequence #status predicted <SIG>
 F:35-1375/Product: glucosyltransferase #status predicted <MAT>

F:1126-1145/Domain: cpl repeat homology <CP1>
F:1253-1272/Domain: cpl repeat homology <CP2>
F:1318-1337/Domain: cpl repeat homology <CP3>

Query Match 63.6%; Score 68; DB 2; Length 1375;
Best Local Similarity 70.0%; Pred. No. 0.0078;
Matches 14; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 3 NYIFRADSEVQTVIAKII 22
Db 580 SYSFIRADSEVQDIIRNII 599

RESULT 12

A:8175
glucosyltransferase precursor - Streptococcus sobrinus

C/Date: 28-Aug-1992 #sequence_revision 28-Aug-1992 #text_change 15-Oct-1999

C/Species: Streptococcus sobrinus

C/Accession: A38175

R:Abdo, H.; Mtsamura, T.; Kodama, T.; Ohta, H.; Fukui, K.; Kato, K.; Kagawa, H.

J. Bacteriol. 173, 989-996, 1991

A/Title: Peptide sequences for sucrose splitting and glucan binding within Streptococcus

A/Reference number: A38175; MUID:91123227; PMID:1704006

A/Accession: A38175

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-1592 <ABO>

A/Cross-references: GB:D90213; NID:G217032; PIDN:BA14241.1; PID:dl014946; PID:g217033

C/Superfamily: cpl repeat homology

C/Accession: F:1093-1112/Domain: cpl repeat homology <CP1>

F:1222-1241/Domain: cpl repeat homology <CP2>

F:1287-1306/Domain: cpl repeat homology <CP3>

F:1330-1351/Domain: cpl repeat homology <CP4>

F:1352-1371/Domain: cpl repeat homology <CP5>

F:1402-1420/Domain: cpl repeat homology <CP6>

F:1465-1484/Domain: cpl repeat homology <CP7>

F:1513-1532/Domain: cpl repeat homology <CP8>

Query Match 59.8%; Score 64; DB 2; Length 1592;
Best Local Similarity 65.0%; Pred. No. 0.043;
Matches 13; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Qy 3 NYIFRADSEVQTVIAKII 22
Db 550 SYSFIRADSEVQDIIRNII 569

RESULT 13

B84213

hypothetical protein eif2a [imported] - Halobacterium sp. NRC-1

C/Species: Halobacterium sp. NRC-1

C/Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 16-Feb-2001

C/Accession: B84213

R:Ng, W.Y.; Kennedy, S.P.; Mahatras, G.G.; Bergquist, B.; Pan, M.; Shukla, H.D.; Lasky, S.

; Leitner, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jablo

Jun, K.H.; Alam, M.; Freitas, T.

Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000

A:Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ehardt, H.; Lowe, T.M.; Li

A/Title: Genome sequence of Halobacterium species NRC-1.

A/Reference number: A84160; MUID:20504483; PMID:11016950

A/Accession: B84213

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-267 <STO>

A/Cross-references: GB:AE004437; NID:gl0580148; PIDN:AA019070.1; GSPDB:GN0C138

C/Genetics:

A:Gene: eif2a

C/Superfamily: translation initiation factor eIF-2 alpha chain

Query Match 42.1%; Score 45; DB 2; Length 267;
Best Local Similarity 50.0%; Pred. No. 8.8;
Matches 8; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

Qy 7 IRADSEVQTVIAKII 22
Db 52 VRDVEDQTVIAKVL 67

RESULT 14

pyruvate kinase (EC 2.7.1.40) A, chloroplast - common tobacco

C/Species: Nicotiana tabacum (common tobacco)

C/Date: 28-Oct-1996 #sequence_revision 07-Feb-1997 #text_change 18-Jun-1999

C/Accession: S51946; S44286

R:Blakeley, S.; Gottlob-McHugh, S.; Wan, J.; Crews, L.; Miki, B.; Ko, K.; Dennis, D.T.

Plant Mol. Biol. 27, 79-89, 1995

A/Title: Molecular characterization of plastid pyruvate kinase from castor and tobacco.

A/Reference number: S51946; MUID:95170010; PMID:7865798

A/Accession: S51946

A>Status: nucleic acid sequence not shown

A:Molecule type: mRNA

A:Residues: 1-593 <BLA>

A/Cross-references: EMBL:Z28373

R:Blakeley, S.; Gottlob-McHugh, S.; Wan, J.; Crews, L.; Miki, B.; Ko, K.; Dennis, D.

submitted to the EMBL Data Library, November 1993

A/Description: Molecular characterization of plastid pyruvate kinase from castor and tob

A/Reference number: S44286

A/Accession: S44286

A:Molecule type: mRNA

A:Residues: 1-86, A', 88-593 <BLM>

A/Cross-references: EMBL:Z28373; NID:G482935; PIDN:CA062222.1; PID:G482935

C/Genetics:

A:Genome: nuclear

C/Superfamily: pyruvate kinase

C/Keywords: chloroplast; glycolysis; phosphotransferase

Query Match 41.6%; Score 44.5; DB 2; Length 593;
Best Local Similarity 52.4%; Pred. No. 26;
Matches 11; Conservative 5; Mismatches 4; Indels 1; Gaps 1;

Qy 1 MANYIFRADSEVQTVIAKII 21
Db 323 LKSYIQARSDSI-SVIAKI 342

RESULT 15

mutator MUTR protein VC2392 [imported] - Vibrio cholerae (strain N16961 serogroup O1)

C/Species: Vibrio cholerae

C/Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Feb-2001

C/Accession: D82080

R:Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwin, M.L.; Dodson, R.J.;

Chadson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoti, I.; Sellers, P

1, R.R.; Mekalanos, J.J.; Venter, J.C.; Frazer, C.M.

Nature 406, 477-483, 2000

A/Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.

A/Reference number: A82035; MUID:20406833; PMID:10952301

A/Accession: D82080

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-132 <HEI>

A/Cross-references: GB:AE004310; GB:AE003852; NID:9656963; PIDN:AA095335.1; GSPDB:GN001

A/Experimental source: serogroup O1; strain N16961; biotype El Tor

C/Genetics:

A:Gene: VC2392

A:Map position: 1

C/Superfamily: mutator MUTR; mutR domain homology

Query Match 41.1%; Score 44; DB 2; Length 132;
Best Local Similarity 50.0%; Pred. No. 5.9;
Matches 10; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

Qy 1 MANYIFRADSEVQTVIAK 20
Db 111 LANYRFEANDPVVKVIAQ 130

Thu Nov 13 12:12:17 2003

us-09-290-049a-16.rpt

Page 5

Search completed: November 13, 2003, 09:50:30
Job time : 15.6588 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 13, 2003, 09:31:40 ; Search time 7.32417 Seconds
(without alignments)
130.561 Million cell updates/sec

Title: US-09-290-049A-16

Perfect score: 107
Sequence: 1 MANYIFRAHDSVQVIANKII 22

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_41:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	107	100.0	1462	1	GTFD_STRMU
2	90	84.1	1365	1	GPR5_STRDO
3	68	63.6	1455	1	GTFD_STRMU
4	67	62.6	1476	1	GTFB_STRMU
5	64	59.8	1592	1	GTF2_STRDO
6	62	57.9	1597	1	GTF1_STRDO
7	45	42.1	267	1	IF2A_HAUNI
8	44.5	41.6	593	1	KRYA_TOBAC
9	44	41.1	479	1	PRL2_ARATH
10	42	39.3	205	1	ADEN_AD88
11	42	39.3	236	1	PHOU_XYLF
12	42	39.3	336	1	RA51_DROME
13	42	39.3	1153	1	NOS_LYMS
14	41.5	38.8	583	1	KRYA_RICCO
15	41	38.3	368	1	RH23_SCHPO
16	41	38.3	411	1	VP48_NVOOP
17	41	38.3	602	1	SYD_TREPA
18	40	37.4	344	1	CYS5_SYMP7
19	40	37.4	417	1	AGP_PRORE
20	40	37.4	557	1	SYD_MYCPN
21	40	37.4	654	1	CBP1_YEAST
22	40	37.4	1018	1	VGNM_BPMV
23	40	37.4	1246	1	SKIM_HUMAN
24	39.5	36.9	231	1	BIOD_YEAST
25	39.5	36.9	332	1	COBS_PSEDE
26	39.5	36.9	1082	1	RROO_RORPC
27	39.5	36.9	146	1	V224_FOMPY
28	39	36.4	238	1	V224_FOMPY
29	39	36.4	238	1	V224_FOMPY
30	39	36.4	272	1	V381_TREPA
31	39	36.4	283	1	LEGI_HAECO
32	39	36.4	297	1	LEC3_CAEEL
33	39	36.4	350	1	VOD2_DROME

34	39	36.4	351	1	RFBB_XANCP	P55295 xanthomonas
35	39	36.4	364	1	PLCE_HUMAN	Q9nuq2 homo sapien
36	39	36.4	380	1	ADH_MALDO	P48977 malus domes
37	39	36.4	433	1	TCR_STAAU	P02983 straphylococ
38	39	36.4	434	1	SYN_PYPUR	Q8u4d3 pyrococcus
39	39	36.4	461	1	GTFB_PLETA	Q9fax2 flexibacter
40	39	36.4	525	1	COX1_CAEEL	P24893 caenorhabdi
41	39	36.4	548	1	LIP2_CANRU	P32946 candida rug
42	39	36.4	549	1	LIP3_CANRU	P32947 candida rug
43	39	36.4	549	1	LIP4_CANRU	P32948 candida rug
44	39	36.4	586	1	SYD_EORPU	Q51402 borrelia bu
45	39	36.4	592	1	SYD_THERTN	Q8rai7 thermocaneer

ALIGNMENTS

RESULT 1	GTFD_STRMU	STANDARD; PRT; 1462 AA.
ID	GTFD_STRMU	AC P49331; 069383; 069386; 069389; 069392; 069398;
DT	01-FEB-1996 (Rel. 33, Created)	
DT	28-FEB-2003 (Rel. 41, Last sequence update)	
DT	15-SEP-2003 (Rel. 42, Last annotation update)	
DE	Glucosyltransferase-S precursor (EC 2.4.1.5) (GTF-S) (Dextranucrase)	
DE	(Sucrose 6-glucosyltransferase).	
GN	GTFD OR SMU.910.	
OS	Streptococcus mutans.	
OC	Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;	
OC	Streptococcus.	
OX	NCBI_TaxID=1309;	
RN	[1]	
RP	SEQUENCE FROM N.A.	
RC	STRAIN=GS-5;	
RC	MEDLINE=91100958; Pubmed=2148600;	
RA	Honda O., Kato C., Kuramitsu H.K.;	
RT	"Nucleotide sequence of the Streptococcus mutans gtfD gene encoding the glucosyltransferase-S enzyme.";	
RL	J. Gen. Microbiol. 136:2099-2105(1990).	
RN	[2]	
RP	SEQUENCE FROM N.A.	
RC	STRAIN=MT4239 / Serotype C, MT4245 / Serotype E, MT4251 / Serotype F, MT4467 / Serotype B, and MT8148 / Serotype C;	
RC	MEDLINE=98231643; Pubmed=9570124;	
RA	Fujiwara T., Terao Y., Hoshino T., Kawabata S., Ooshima T., Sobue S., Kimura S., Hamada S.;	
RT	"Molecular analyses of glucosyltransferase genes among strains of Streptococcus mutans.";	
RT	FEMS Microbiol. Lett. 161:331-336(1998).	
RN	[3]	
RP	SEQUENCE FROM N.A.	
RC	STRAIN=UA159 / Serotype C;	
RC	MEDLINE=2295063; Pubmed=12397186;	
RA	Aldic D., McShan W.M., McLaughlin R.E., Savic G., Chang J., Carson W.B., Primeaux C., Tian R., Kenton S., Jia H., Lin S., Qian Y., Li S., Zhu H., Najaf F., Lai B., White J., Roe B.A., Perrecci J.O.;	
RT	"Genome sequence of Streptococcus mutans UA159, a cariogenic dental pathogen.";	
RT	Proc. Natl. Acad. Sci. U.S.A. 99:14434-14439(2002).	
CC	-1- FUNCTION: PRODUCTION OF EXTRACELLULAR GLUCANS, THAT ARE THOUGHT TO PLAY A KEY ROLE IN THE DEVELOPMENT OF THE DENTAL PLAQUE BECAUSE OF THEIR ABILITY TO ADHERE TO SMOOTH SURFACES AND MEDIATE THE AGGREGATION OF BACTERIAL CELLS AND FOOD DEBRIS.	
CC	-1- CATALYTIC ACTIVITY: Sucrose + ((1,6)-alpha-D-glucosyl) (N) = D-fructose + ((1,6)-alpha-D-glucose).	
CC	-1- SUBCELLULAR LOCATION: Secreted.	
CC	-1- DISEASE: DENTAL CARIES.	
CC	-1- MISCELLANEOUS: GTF-I SYNTHESIZES WATER-INSOLUBLE GLUCANS (ALPHA 1,3-LINKED GLUCOSE AND SOME 1,6 LINKAGES), GTF-S SYNTHESIZES WATER-SOLUBLE GLUCANS (ALPHA 1,6-GLUCOSE). GTF-SI SYNTHESIZES BOTH FORMS OF GLUCANS.	
CC	-1- SIMILARITY: TO OTHER GLUCOSYLTRANSFERASES AND SOME TO A GLUCAN-BINDING PROTEIN FROM S. MUTANS.	


```

CC -1- SIMILARITY: Contains 6 cell wall binding repeats.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M29236; AAA26895.1; -
DR EMBL; D88653; BAA26103.1; -
DR EMBL; D88656; BAA26107.1; -
DR EMBL; D88659; BAA26111.1; -
DR EMBL; D88662; BAA26115.1; -
DR EMBL; D89979; BAA26121.1; -
DR EMBL; AE014932; AAN58619.1; -
DR InterPro; IPR002479; CW_binding_70.
DR InterPro; IPR003318; Glyco_hydro_70.
DR Pfam; PF02324; Glyco_hydro_70; 1.
KM Complete proteome.
KW Complete proteome.
FT CHAIN 1 1462
FT DOMAIN 1232 1423
FT REPEAT 1232 1295
FT REPEAT 1296 1359
FT REPEAT 1360 1423
FT REPEAT 10 10
FT VARIANT 19 19
FT VARIANT 58 58
FT VARIANT 68 68
FT VARIANT 81 81
FT VARIANT 113 113
FT VARIANT 122 122
FT VARIANT 132 132
FT VARIANT 135 135
FT VARIANT 137 137
FT VARIANT 202 202
FT VARIANT 255 255
FT VARIANT 275 275
FT VARIANT 288 288
FT VARIANT 301 301
FT VARIANT 313 313
FT VARIANT 317 317
FT VARIANT 328 328
FT VARIANT 350 350
FT VARIANT 628 628
FT VARIANT 688 688
FT VARIANT 726 730
FT VARIANT 762 762
FT VARIANT 964 964
FT VARIANT 1019 1019
FT VARIANT 1059 1060
FT VARIANT 1060 1060
FT VARIANT 1080 1080
FT VARIANT 1142 1142
FT VARIANT 1198 1198
FT VARIANT 1220 1220
FT VARIANT 1280 1280
FT VARIANT 1282 1282
FT VARIANT 1290 1290
FT VARIANT 1311 1311

```

```

FT VARIANT 1403 1403 G -> D (IN STRAINS GS-5 AND MT4467).
FT VARIANT 1425 1425 G -> R (IN STRAIN GS-5).
FT VARIANT 1449 1449 R -> K (IN STRAIN MT4467).
FT CONFLICT 1428 1462 RYDKNSGMYYNKKVTLANRRIGIDRWGIARY -> VY
R (IN REF. 1)
SQ SEQUENCE 1462 AA; 163387 MW; CE4A279C4D708645 CRC64;
Query Match 100.0%; Score 107; DB 1; Length 1462;
Best Local Similarity 100.0%; Pred. No. 7.1e-10;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MANYFIPIAHDSVQTVIAKII 22
DB 574 MANYFIPIAHDSVQTVIAKII 595
RESULT 2
GTF5_STRDO STANDARD; PRT; 1365 AA.
AC P29336;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Glucosyltransferase-S precursor (EC 2.4.1.5) (GTF-S) (Dextranucrase)
DE (Sucrose 6-glucosyltransferase).
GTF5.
GN Streptococcus downei (Streptococcus sobrinus).
OS Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OC NCBI_TaxID=1317;
RN (1)
RP SEQUENCE FROM N.A.
RX MEDLINE=90316665; PubMed=2142479;
RA Gilmore K.S., Russell R.R., Ferretti J.J.;
RT "Analysis of the Streptococcus downei gtf5 gene, which specifies a
RT glucosyltransferase that synthesizes soluble glucans.";
RL Infect. Immun. 58:2452-2458 (1990).
CC -1- FUNCTION: PRODUCTION OF EXTRACELLULAR GLUCANS, THAT ARE THOUGHT
CC TO PLAY A KEY ROLE IN THE DEVELOPMENT OF THE DENTAL PLAQUE BECAUSE
CC OF THEIR ABILITY TO ADHERE TO SMOOTH SURFACES AND MEDIANE THE
CC AGGREGATION OF BACTERIAL CELLS AND FOOD DEBRIS.
CC -1- CATALYTIC ACTIVITY: Sucrose + ((1,6)-alpha-D-glucosyl) (N) = D-
CC fructose + ((1,6)-alpha-D-glucosyl) (N+1).
CC -1- ENZYME REGULATION: GLUCAN SYNTHESIS BY GTF-S IS INDEPENDENT OF
CC PRIMER GLUCAN UNLIKE GTF-I.
CC -1- DISEASE: DENTAL CARIES.
CC -1- MISCELLANEOUS: GTF-S SYNTHESIZES WATER-SOLUBLE GLUCANS (ALPHA
CC 1,6-GLUCOSE).
CC -1- SIMILARITY: TO OTHER GLUCOSYLTRANSFERASES AND SOME TO A GLUCAN-
CC BINDING PROTEIN FROM S. MUTANS.
CC -1- SIMILARITY: Contains 10 cell wall binding repeats.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M30943; AAA26898.1; -
DR InterPro; IPR002479; CW_binding.
DR InterPro; IPR003318; Glyco_hydro_70.
DR Pfam; PF01473; CW_binding_1; 8.
DR Pfam; PF02324; Glyco_hydro_70; 1.
KW Transferase; Glucosyltransferase; Signal; Repeat; Dental caries.
FT SIGNAL 1 36
FT CHAIN 37 1365
FT REPEAT 157 177
FT REPEAT 178 197
FT DOMAIN 198 1061
FT REPEAT 1062 1082

```

FT REPEAT 1083 1102 CELL WALL BINDING 4.
 FT REPEAT 1150 1169 CELL WALL BINDING 5.
 FT REPEAT 1170 1190 CELL WALL BINDING 6.
 FT REPEAT 1225 1243 CELL WALL BINDING 7.
 FT REPEAT 1289 1308 CELL WALL BINDING 8.
 FT REPEAT 1309 1328 CELL WALL BINDING 9.
 FT REPEAT 1331 1352 CELL WALL BINDING 10.
 SQ SEQUENCE 1365 AA; 151590 MW; 1672965A2E8C476 CRC64;
 Query Match 84.1%; Score 90; DB 1; Length 1365;
 Best Local Similarity 90.0%; Pred. No. 5e-07;
 Matches 18; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 Oy 3 NYIFRAHDSVQVIAKII 22
 Db 539 NYVIFRAHDSVQVIAKII 558
 RESULT 3
 GTF-STRMU STANDARD; PRT; 1455 AA.
 ID GTF-STRMU PRT; 1455 AA.
 AC P13470; 069382; 069388; 069391; 069397; P05427;
 DT 01-NOV-1998 (Rel. 09, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update) (GTF-SI)
 DE Glucosyltransferase-SI precursor (EC 2.4.1.5) (GTF-SI)
 DE (Dextrantransferase) (Sucrose 6-glucosyltransferase).
 GN GTF-OR SMU.1005.
 OS Streptococcus mutans.
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 CX NCBI_TaxID=1309;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=GS-5;
 RX MEDLINE=89137980; PubMed=2976010;
 RA Ueda S., Shiroza T., Kuramitsu H.K.;
 RT "Sequence analysis of the gtfC gene from Streptococcus mutans GS-5";
 RL Gene 69:101-109(1988).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MT4239 / Serotype C, MT4245 / Serotype E, MT4251 / Serotype F,
 MT4467 / Serotype E, and MT8148 / Serotype C;
 RX MEDLINE=98231643; PubMed=9570124;
 RA Fujiwara T., Terao Y., Hoshino T., Kawabata S., Oshima T., Sobue S.,
 RA Kimura S., Hamada S.;
 RT "Molecular analyses of glucosyltransferase genes among strains of
 RT Streptococcus mutans";
 RL FEMS Microbiol. Lett. 161:331-336(1998).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=UA159 / ATCC 700610 / Serotype C;
 RX MEDLINE=22295063; PubMed=12397186;
 RA Ajdic D., McMan W.M., McLaughlin R.E., Savic G., Chang J.,
 RA Carson M.B., Primeaux C., Tian R., Kenton S., Jia H., Jin S., Qian Y.,
 RA Li S., Zhu H., Najjar H., Lai H., White J., Roe B.A., Perretti J.J.;
 RT "Genome sequence of Streptococcus mutans UA159, a cariogenic dental
 RT pathogen";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:14434-14439(2002).
 RN [4]
 RP SEQUENCE OF 1-349 FROM N.A.
 RC STRAIN=GS-5;
 RX MEDLINE=87308013; PubMed=3040685;
 RA Shiroza T., Ueda S., Kuramitsu H.K.;
 RT "Sequence analysis of the gtfC gene from Streptococcus mutans";
 RL J. Bacteriol. 169:4263-4270(1987).
 CC -1- FUNCTION: PRODUCTION OF EXTRACELLULAR GLUCANS THAT ARE THOUGHT
 CC TO PLAY A KEY ROLE IN THE DEVELOPMENT OF THE DENTAL PLAQUE BECAUSE
 CC OF THEIR ABILITY TO ADHERE TO SMOOTH SURFACES AND MEDIANE THE
 CC AGGREGATION OF BACTERIAL CELLS AND FOOD DEBRIS.
 CC -1- CATALYTIC ACTIVITY: Sucrose + (1,6)-alpha-D-glucosyl(1) (N) = D-
 CC fucose + (1,6)-alpha-D-glucosyl(1) (N+1).
 CC -1- SUBCELLULAR LOCATION: Secreted.

CC -1- DISEASE: DENTAL CARIES.
 CC -1- MISCELLANEOUS: GTF-I SYNTHESIZES WATER-INSOLUBLE GLUCANS (ALPHA
 CC 1,3-LINKED GLUCOSE AND SOME 1,6 LINKAGES), GTF-S SYNTHESIZES
 CC WATER-SOLUBLE GLUCANS (ALPHA 1,6-GLUCOSE). GTF-SI SYNTHESIZES BOTH
 CC FORMS OF GLUCANS.
 CC -1- SIMILARITY: TO OTHER GLUCOSYLTRANSFERASES AND SOME TO A GLUCAN-
 CC BINDING PROTEIN FROM S. MUTANS.
 CC -1- SIMILARITY: Contains 5 cell wall binding repeats.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.ebi.ac.uk/announcements/>
 CC or send an email to license@ebi.ac.uk).
 CC -----
 DR EMBL; M22054; AAA88592.1; -;
 DR EMBL; D88652; BAA26102.1; -;
 DR EMBL; D88655; BAA26106.1; -;
 DR EMBL; D88658; BAA26110.1; -;
 DR EMBL; D88661; BAA26114.1; -;
 DR EMBL; D89978; BAA26120.1; -;
 DR EMBL; A8014840; AAN58706.1; -;
 DR EMBL; M17361; AAA88589.1; -;
 DR InterPro; IPR002479; CW binding.
 DR InterPro; IPR003318; Glyco hydro 70.
 DR Pfam; PF01473; CW binding 1; 1.
 DR Pfam; PF02324; Glyco hydro 70; 1.
 KW Transferase; Glucosyltransferase; Signal; Repeat; Dental caries;
 KW Complete proteome.
 FT SIGNAL 1 34
 FT CHAIN 35 1455
 FT DOMAIN 35 1050
 FT DOMAIN 1126 1455
 FT DOMAIN 1126 1455
 FT REPEAT 1126 1159
 FT REPEAT 1169 1200
 FT REPEAT 1227 1238
 FT REPEAT 1253 1303
 FT REPEAT 1318 1330
 FT REPEAT 21 21
 FT REPEAT 81 81
 FT REPEAT 106 106
 FT REPEAT 116 116
 FT REPEAT 126 126
 FT REPEAT 150 151
 FT VARIANT 256 256
 FT VARIANT 425 425
 FT VARIANT 519 519
 FT VARIANT 538 538
 FT VARIANT 545 545
 FT VARIANT 597 597
 FT VARIANT 600 600
 FT VARIANT 601 601
 FT VARIANT 614 614
 FT VARIANT 727 727
 FT VARIANT 734 734
 FT VARIANT 964 964
 FT VARIANT 1113 1113
 FT VARIANT 1118 1118
 FT VARIANT 1204 1204
 FT VARIANT 1208 1208
 FT VARIANT 1292 1294
 FT VARIANT 1305 1369
 FT VARIANT 1326 1326
 FT VARIANT 1331 1331
 A -> V (IN STRAINS GS-5 AND MT4467).
 A -> T (IN STRAIN GS-5).
 M -> I (IN STRAIN GS-5).
 T -> I (IN STRAIN MT8148).
 A -> V (IN STRAIN MT8148).
 L -> F (IN STRAIN MT4239).
 N -> Y (IN STRAIN MT4239).
 I -> T (IN STRAIN MT4239).
 A -> V (IN STRAINS GS-5, MT4239, MT4467
 AND MT8148).
 V -> I (IN STRAIN MT8148).
 DGH -> NGY (IN STRAINS GS-5, MT4467 AND
 MT8148).
 MISSING (IN STRAIN MT4245).
 I -> V (IN STRAIN MT8148).
 T -> A (IN STRAINS GS-5, MT4239, MT4467
 AND MT8148).

```
FT VARIANT 1377 1377 R -> K (IN STRAIN MT8148).
FT VARIANT 1398 1398 V -> I (IN STRAIN MT8148).
FT VARIANT 1424 1424 D -> N (IN STRAIN MT4239).
FT VARIANT 1439 1439 V -> I (IN STRAINS MT4239 AND MT8148).
FT VARIANT 1444 1444 S -> P (IN STRAIN MT8148).
FT CONFLICT 1337 1455 ORLYPKNSGVQAKGELLTERGRIKIVYPPNSGNEVNRNRYR
TSSGMYEYNGDVALIGMHVVEGRVRYFDNGYRYASHD
ORNMWDYDRRDPCGGSSAVRFRHSRNGFDNFRF ->
HASILSLMVFRLRESLSQSVKVSNTMLIFEMKFLVIM
(IN REF. 1).
SQ SEQUENCE 1455 AA; 162965 MW; 3CB455A994FEC86 CRC64;

Query Match 63.6%; Score 68; DB 1; Length 1455;
Beat Local Similarity 70.0%; Pred. No. 0.0029;
Matches 14; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Oy 3 NYIFIRAHSEVQTVIAKII 22
Db 580 SYSFIRAHSEVQDIIRNII 599

RESULT 4
GTFB_STRMU STANDARD; PRT; 1476 AA.
AC P08987; 069381; 069387; 069390; 069396;
DT 01-NOV-1988 (Rel. 09; Created)
DT 28-SEP-2003 (Rel. 41; Last sequence update)
DT 15-SEP-2003 (Rel. 42; Last annotation update)
DE Glucosyltransferase-1 precursor (EC 2.4.1.5) (GTF-I) (Dextranucrase)
GN (Sucrose 6-glucosyltransferase).
OS GTFB OR SMU.1004.
OC Streptococcus mutans.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1309;
[1]
RN SEQUENCE FROM N.A.
RP STRAIN-GS-5;
RC MEDLINE=87308013; PubMed=3040685;
RA Shirota T., Ueda S., Kuramitsu H.K.;
RT "Sequence analysis of the gtfb gene from Streptococcus mutans.";
RL J. Bacteriol. 169:4263-4270(1987).
[2]
RN SEQUENCE FROM N.A.
RP STRAIN-MT4239 / Serotype C, MT4245 / Serotype E, MT4251 / Serotype F,
RC MT4467 / Serotype E, and MT8148 / Serotype C;
RA MEDLINE=98231643; PubMed=9570124;
RA Fujiwara T., Terao Y., Hoshino T., Kawabata S., Ooshima T., Sobue S.,
RA Kimura S., Hamada S.;
RT "Molecular analyses of glucosyltransferase genes among strains of
RT Streptococcus mutans.";
RL FEMS Microbiol. Lett. 161:331-336(1998).
[3]
RN SEQUENCE FROM N.A.
RP STRAIN-UA159 / ATCC 700610 / Serotype C;
RC STRAIN-UA159 / ATCC 700610 / Serotype C;
RA MEDLINE=22295063; PubMed=12397186;
RA Ajdic D., McShawn W.M., McLoughlin R.E., Savic G., Chang J.,
RA Carson M.B., Prineaux C., Tian R., Kenton S., Jia H., Lin S., Qian Y.,
RA Li S., Zhu H., Najjar F., Lai H., White J., Roe B.A., Perrett J.U.;
RT "Genome sequence of Streptococcus mutans UA159, a cariogenic dental
RT pathogen.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:14434-14439(2002).
CC -1- FUNCTION: PRODUCTION OF EXTRACELLULAR GLUCANS, THAT ARE THOUGHT
CC TO PLAY A KEY ROLE IN THE DEVELOPMENT OF THE DENTAL PLAQUE BECAUSE
CC OF THEIR ABILITY TO ADHERE TO SMOOTH SURFACES AND MEDIANE THE
CC AGGREGATION OF BACTERIAL CELLS AND FOOD DEBRIS.
CC -1- CATALYTIC ACTIVITY: Sucrose + {(1,6)-alpha-D-glucosyl} (N) = D-
CC fructose + {(1,6)-alpha-D-glucosyl} (N+1).
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- DISEASE: DENTAL CARIES.
CC -1- MISCELLANEOUS: GTF-I SYNTHESIZES WATER-INSOLUBLE GLUCANS (ALPHA
CC 1,3-LINKED GLUCOSE AND SOME 1,6 LINKAGES), GTF-S SYNTHESIZES
CC WATER-SOLUBLE GLUCANS (ALPHA 1,6-GLUCOSE). GTF-SI SYNTHESIZES BOTH
CC
```

```

FT. VARIANT 1329 1329 H -> Y (IN STRAINS GS-5 AND MT4467).
FT. VARIANT 1394 1394 Y -> H (IN STRAINS GS-5, MT4239, MT4467
FT. VARIANT 1402 1402 AND MT8148).
FT. VARIANT 1402 1402 S -> G (IN STRAINS GS-5, MT4239, MT4467
FT. VARIANT 1459 1459 AND MT8148).
FT. VARIANT 1459 1459 Y -> H (IN STRAIN MT4467).
FT. CONFLICT 570 570 R -> A (IN REF. 1).
FT. CONFLICT 800 817 ADDVRVAASTABSTDSK -> LIRKMLALRPHQOMA
FT. CONFLICT 800 817 (IN REF. 1).
FT. CONFLICT 1310 1310 H -> L (IN REF. 1).
SQ. SEQUENCE 1476 AA; 165846 MW; 9C6E09F731B4CBFC CRC64;

Query Match 62.6%; Score 67; DB 1; Length 1476;
Best Local Similarity 70.0%; Pred. No. 0.0043;
Matches 14; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 3 NYIFIRAHSEVQVTIAKII 22
Db 554 SYSFIRAHSEVQDILRDII 573

RESULT 5
GTF2_STRDO STANDARD; PRT; 1592 AA.
AC P27470;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Glucosyltransferase-I precursor (EC 2.4.1.5) (GTF-I) (Dextranucrase)
DE (Sucrose 6-glucosyltransferase).
OS Streptococcus downei (Streptococcus sobrinus).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=13117;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=6715 / Serotype G;
RX MEDLINE=9112327; PubMed=1704006;
RA Abo H.; Matsumura T.; Kodama T.; Ohba H.; Kato K.;
RA Kagawa H.;
RT "Peptide sequences for sucrose splitting and glucan binding within
RT Streptococcus sobrinus glucosyltransferase (water-insoluble glucan
RT synthetase)";
RL J. Bacteriol. 173:989-996(1991).
CC -1- FUNCTION: PRODUCTION OF EXTRACELLULAR GLUCANS, THAT ARE THOUGHT
CC TO PLAY A KEY ROLE IN THE DEVELOPMENT OF THE DENTAL PLAQUE BECAUSE
CC OF THEIR ABILITY TO ADHERE TO SMOOTH SURFACES AND MEDIATE THE
CC AGGREGATION OF BACTERIAL CELLS AND FOOD DEBRIS.
CC -1- CATALYTIC ACTIVITY: Sucrose + ((1,6)-alpha-D-glucosyl)(N) = D-
CC fructose + ((1,6)-alpha-D-glucosyl)(N+1).
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- DISEASE: DENTAL CARIES.
CC -1- MISCELLANEOUS: GTF-I SYNTHESIZES WATER-INSOLUBLE GLUCANS (ALPHA
CC 1,3-LINKED GLUCOSE AND SOME 1,6 LINKAGES). GTF-S SYNTHESIZES
CC WATER-SOLUBLE GLUCANS (ALPHA 1,6-GLUCOSE). GTF-SI SYNTHESIZES BOTH
CC FORMS OF GLUCANS.
CC -1- SIMILARITY: TO OTHER GLUCOSYLTRANSFERASES AND SOME TO A GLUCAN-
CC BINDING PROTEIN FROM S. MUTANS.
CC -1- SIMILARITY: Contains 16 cell wall binding repeats.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; D90213; BAAL4241.1; -
CC InterPro; IPR002479; CM binding.
CC InterPro; IPR003318; Glyco_hydro_70.
CC Pfam; PF01473; CM binding_I; 13
CC Pfam; PF02324; Glyco_hydro_70; 1.

```

```

KM Transferase; Glucosyltransferase; Signal; Repeat; Dental caries.
FT SIGNAL 1 38 POTENTIAL.
FT CHAIN 39 1592 GLUCOSYLTRANSFERASE-I.
FT DOMAIN 39 1044 CATALYTIC (APPROXIMATE).
FT DOMAIN 1093 1592 GLUCAN-BINDING (APPROXIMATE)..
FT DOMAIN 1093 1592 6.5 X TANDEM REPEATS.
FT REPEAT 1093 1142 1.
FT REPEAT 1158 1207 2.
FT REPEAT 1222 1272 3.
FT REPEAT 1287 1337 4.
FT REPEAT 1402 1451 5.
FT REPEAT 1514 1563 6.
FT REPEAT 1577 1592 7 (INCOMPLETE).
SQ. SEQUENCE 1592 AA; 176167 MW; BC0A66D079351BCF CRC64;

Query Match 59.8%; Score 64; DB 1; Length 1592;
Best Local Similarity 65.0%; Pred. No. 0.015;
Matches 13; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Qy 3 NYIFIRAHSEVQVTIAKII 22
Db 550 SYSFIRAHSEVQDILRDII 569

RESULT 6
GTF1_STRDO STANDARD; PRT; 1597 AA.
AC P11001;
ID GTF1_STRDO
DT 01-JUN-1989 (Rel. 11, Created)
DT 01-JUN-1989 (Rel. 11, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Glucosyltransferase-I precursor (EC 2.4.1.5) (GTF-I) (Dextranucrase)
DE (Sucrose 6-glucosyltransferase).
OS Streptococcus downei (Streptococcus sobrinus).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=13117;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MFE28;
RX MEDLINE=87308014; PubMed=3040686;
RA Reretti J.J.; Gilpin M.L.; Russell R.R.B.;
RT "Nucleotide sequence of a glucosyltransferase gene from Streptococcus
RT sobrinus MFE28.";
RL J. Bacteriol. 169:4271-4278(1987).
CC -1- FUNCTION: PRODUCTION OF EXTRACELLULAR GLUCANS, THAT ARE THOUGHT
CC TO PLAY A KEY ROLE IN THE DEVELOPMENT OF THE DENTAL PLAQUE BECAUSE
CC OF THEIR ABILITY TO ADHERE TO SMOOTH SURFACES AND MEDIATE THE
CC AGGREGATION OF BACTERIAL CELLS AND FOOD DEBRIS.
CC -1- CATALYTIC ACTIVITY: Sucrose + ((1,6)-alpha-D-glucosyl)(N) = D-
CC fructose + ((1,6)-alpha-D-glucosyl)(N+1).
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- DISEASE: DENTAL CARIES.
CC -1- MISCELLANEOUS: GTF-I SYNTHESIZES WATER-INSOLUBLE GLUCANS (ALPHA
CC 1,3-LINKED GLUCOSE AND SOME 1,6 LINKAGES). GTF-S SYNTHESIZES
CC WATER-SOLUBLE GLUCANS (ALPHA 1,6-GLUCOSE). GTF-SI SYNTHESIZES BOTH
CC FORMS OF GLUCANS.
CC -1- SIMILARITY: TO OTHER GLUCOSYLTRANSFERASES AND SOME TO A GLUCAN-
CC BINDING PROTEIN FROM S. MUTANS.
CC -1- SIMILARITY: Contains 19 cell wall binding repeats.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M17391; AAC63063.1; -
CC InterPro; IPR002479; CM binding.
CC InterPro; IPR003318; Glyco_hydro_70.

```



```

FT ACT SITE 341 341 BY SIMILARITY.
FT METAL 343 343 MAGNESIUM (POTENTIAL).
FT METAL 364 364 MAGNESIUM (POTENTIAL).
FT METAL 365 365 MAGNESIUM (POTENTIAL).
SQ SEQUENCE 593 AA; 65227 MW; 5AF049E193C1D484 CRC64;

Query Match 41.6%; Score 44.5; DB 1; Length 593;
Best Local Similarity 52.4%; Pred. No. 11;
Matches 11; Conservative 5; Mismatches 4; Indels 1; Gaps 1;

Qy 1 MANYIFRADSEVQYTIK 21
Db 323 LKSYIQARDSDI-SYIYAKI 342

RESULT 9
PRL2 ARATH STANDARD; PRT; 479 AA.
ID PRL2 ARATH 094007; Q9LUR9;
AC 039150; 094007; Q9LUR9;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE PPI/PP2A phosphatases pleiotropic regulator PRL2.
OS PRL2 OR AT3G1650 OR MGL6.11.
GN Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC Eurosid II; Brassicales; Brassicaceae; Arabidopsis.
ON NCBI_TaxID=3702;
RX STRAIN=cv. Columbia;
RA Nemeth K., Salchert K., Putnocky P., Bhalerao R., Koncz-Kalman Z.,
RA Stankovic-Stangeland B., Bako L., Mathur J., Okiesz L., Stabel S.,
RA Geisenberger P., Sitt M., Redel G.P., Schnell J., Koncz C.;
RT "pleiotropic control of glucose and hormone responses by PRL1, a
RT nuclear WD protein, in Arabidopsis."
RL Genes Dev. 12:3059-3073(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RX MEDLINE=20277480; PubMed=10819329;
RA Sato S., Nakamura Y., Kaneko T., Katoh T., Asamizu E., Tabata S.;
RT "Structural analysis of Arabidopsis thaliana chromosome 3. I. Sequence
RT features of the regions of 4,504,864 bp covered by sixty P1 and PAC
RT clones."
RL DNA Res. 7:131-135(2000).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RA Shinozaki K., Davis R.W., Ecker J.R., Theologis A.;
RT "RIKEN Arabidopsis full length cDNA clones (RAFLs) sequenced by the
RT SSP consortium (Salk/Stanford/PGEC).";
RL submitted (Aug-2001) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: PLEIOTROPIC REGULATOR OF PPI AND PP2A PHOSPHATASES.
CC -1- SIMILARITY: Contains 7 WD repeats.
CC -1- SIMILARITY: BELONGS TO THE PRL1/PRL2 FAMILY OF WD-REPEAT PROTEINS.
CC -1- CAUTION: Ref.2 sequence differs from that shown due to erroneous
CC gene model prediction.
CC -1- CAUTION: It is uncertain whether Met-1 or Met-3 is the initiator.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X82826; CAA58033.1; -
CC EMBL; AB022217; BAB02756.1; ALT SEQ.
CC EMBL; AY054181; AAL06842.1; ALT_INIT.

```

```

DR InterPro; IPR001680; WD40.
DR Pfam; PF00400; WD40; 7.
DR PRINTS; PR00320; GPOTENINRPT.
DR PRODOM; PD000018; WD40; 2.
DR SMART; SM00320; WD40; 7.
DR PROSITE; PS00678; WD_REPEATS_1; 2.
DR PROSITE; PS50082; WD_REPEATS_2; 4.
DR PROSITE; PS50294; WD_REPEATS_REGION; 1.
KM Repeat, WD repeat.
FT REPEAT 168 198 WD 1.
FT REPEAT 210 240 WD 2.
FT REPEAT 252 282 WD 3.
FT REPEAT 293 323 WD 4.
FT REPEAT 335 364 WD 5.
FT REPEAT 377 406 WD 6.
FT REPEAT 426 456 WD 7.
SQ SEQUENCE 479 AA; 53568 MW; 3019864AAE80670 CRC64;

Query Match 41.1%; Score 44; DB 1; Length 479;
Best Local Similarity 50.0%; Pred. No. 10;
Matches 8; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

Qy 5 IFRPADSEVQYTIK 20
Db 289 IFVLPHSDVFSYLAR 304

RESULT 10
ADEN_ADEG8 STANDARD; PRT; 205 AA.
ID ADEN_ADEG8 09QW72;
AC 09QW72;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Adenin (EC 3.4.22.39) (Endoprotease) (Late L3 23 kDa protein).
OS Avian adenovirus type 8 (Strain ATCC A-24) (Fowl adenovirus 8).
OC Viruses; dsDNA viruses, no RNA stage; Adenoviridae; Aviadenovirus.
ON NCBI_TaxID=6295;
RX [1]
RP SEQUENCE FROM N.A.
RA Ojic D., Nagy E.;
RT "The DNA sequence of fowl adenovirus 8."
RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: THIOI PROTEASE CLEAVING SPECIFIC GLY-ALA PEPTIDES IN A
CC NUMBER OF VIRAL PRECURSOR PROTEINS (TERMINAL, IIA, VI, VII, VIII,
CC IIA AND MU) WHICH ARE REQUIRED FOR VIRION MATURATION. ALSO CLEAVES
CC HOST CELLS CYTOSKELETAL KERATINS K7 AND K18.
CC -1- CATALYTIC ACTIVITY: Cleaves proteins of the adenovirus and its
CC host cell at two consensus sites: -Yaa-Xaa-Gly-Gly- and Xaa
CC -Yaa-Xaa-Gly-Xaa-|-Gly- (in which yaa is Met, Ile or Leu, and Xaa
CC is any amino acid).
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY C5.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF083975; AAD50345.2; -
CC HSRP; P03252; IAVP.
CC MEROPS; C05.001; -.
DR InterPro; IPR000855; Peptidase_C5.
DR Pfam; PF00770; Peptidase_C5; 1.
DR PRINTS; PR00703; ADENOPPTASE.
DR PRODOM; PD003705; Peptidase_C5; 1.
DR Hydrolase; Thiol protease; Late protein.
KM ACT SITE 55 55 BY SIMILARITY.
FT ACT_SITE 72 72 BY SIMILARITY.
FT ACT_SITE 122 122 BY SIMILARITY.
SQ SEQUENCE 205 AA; 23701 MW; 36F0700CDFB85F62 CRC64;

```

Query Match 39.3%; Score 42; DB 1; Length 205;
 Best Local Similarity 53.3%; Pred. No. 9;
 Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Qy 1 MANYIFIRADSEVO 15
 Db 179 MKNSFPRAHSEIK 193

RESULT 11

PHOU_XYLFA STANDARD; PRT; 236 AA.
 AC Q9PB09; 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Phosphate transport system protein phou.
 GN PHOU OR XP2145.
 OS Xylella fastidiosa.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
 CC Xanthomonadaceae; Xylella.
 OX NCBI_TaxID=23711;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=9A5C;
 RX MEDLINE=20365717; PubMed=10910347;

RA Simpson A.J.G., Reinach F.C., Arruda P., Abreu F.A., Acencio M.,
 RA Alvaranga R., Alves L.M.C., Araya J.B., Baia G.S., Baptista C.S.,
 RA Barros M.H., Bonaccorsi E.D., Bordin S., Bove U.M., Briones M.R.S.,
 RA Bueno M.R.P., Camargo A.A., Camargo L.E.A., Carraro D.M., Carter H.,
 RA Colauto N.B., Colombo C., Costa F.F., Costa M.C.R., Costa-Neto C.M.,
 RA Coutinho L.L., Cristofani M., Dias-Neto E., Docena C., El-Dorri H.,
 RA Faccinoni A.P., Ferreira A.J.S., Ferreira V.C.A., Ferro J.A.,
 RA Fraga J.S., Franca S.C., Franco M.C., Frohne M., Furian L.R.,
 RA Garner M., Goldman G.H., Goldman M.H.S., Gomes S.L., Graber A.,
 RA Ho P.L., Hohenseil J.D., Junqueira M.L., Kemper E.D., Kitajima J.P.,
 RA Krieger J.E., Kuramae E.E., Laigret F., Lambais M.R., Leite L.C.C.,
 RA Lemos E.G.M., Lemos M.V.F., Lopes S.A., Lopes C.R., Machado J.A.,
 RA Machado W.A., Madeira A.M.B.N., Madeira H.M.F., Mariano C.L.,
 RA Marques M.V., Martins E.A.L., Martins E.M.F., Matsukuma A.Y.,
 RA Menck C.F.M., Miranca E.C., Miyaki C.Y., Monteiro-Vitorino C.B.,
 RA Moon D.H., Nagai M.A., Nascimento A.L.T.O., Netto L.E.S.,
 RA Nham A.Jr., Nobrega F.G., Nunes L.R., Oliveira M.A.,
 RA de Oliveira M.C., de Oliveira R.C., Palmieri D.A., Paris A.,
 RA Peixoto B.R., Pereira G.A.G., Pereira H.A. Jr., Peguero J.B.,
 RA Quaggio R.B., Roberto P.G., Rodrigues V., de Rosa A.U.M.,
 RA da Rosa V.E. Jr., de Sa R.G., Santelli R.V., Sawasaki H.B.,
 RA da Silva A.C.R., da Silva A.M., da Silva F.R., Silva W.A. Jr.,
 RA de Souza A.P., Terenzi M.F., Truffi D., Tsai S.M., Tshako M.H.,
 RA Vallada H., Van Sluys M.A., Verjovski-Almeida S., Vettore A.L.,
 RA Zago M.A., Zatz M., Weidman U., Setubal J.C.;
 RT "The genome sequence of the plant pathogen *Xylella fastidiosa*."

RL Nature 406:151-159(2000).
 CC -1- FUNCTION: NOT KNOWN; PROBABLY INVOLVED IN PHOSPHATE TRANSPORT
 AND/OR METABOLISM (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -1- SIMILARITY: BELONGS TO THE PHOU FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL; AE004029; AAF84944.1; -;
 DR PIR; E82593; E82593.
 DR InterPro; IPR002754; Phou.
 DR Pfam; PF01895; Phou; 2. Complete proteome.
 KW Phosphate transport; Complete proteome.

SQ SEQUENCE 236 AA; 26596 MW; 2B0FB29D4A2F14D1 CRC64;
 Query Match 39.3%; Score 42; DB 1; Length 236;
 Best Local Similarity 30.0%; Pred. No. 10;
 Matches 6; Conservative 8; Mismatches 6; Indels 0; Gaps 0;

Qy 3 NYIFIRADSEVOYTIKII 22
 Db 6 NYHIVKSYDDENRLVTEII 25

RESULT 12

RA51_DROME STANDARD; PRT; 336 AA.
 ID RA51_DROME
 AC Q27297; Q9VAA8; 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE DNA repair protein Rad51 homolog (RecA protein homolog).
 GN RAD51 OR DMR OR CG7948.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 CC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Canton-S;
 RX MEDLINE=95161094; PubMed=7857671;
 RA Akaboshi E., Inoue Y., Iryo H.;
 RT "Cloning of the cDNA and genomic DNA that correspond to the recA-like
 RT gene of *Drosophila melanogaster*."
 RL Jpn. J. Genet. 69:663-670(1994).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Oregon-R;
 RX MEDLINE=96207535; PubMed=8625736;
 RA McKee B.D., Ren X.T., Hong C.S.;
 RT "A recA-like gene in *Drosophila melanogaster* that is expressed at
 RT high levels in female but not male meiotic tissues."
 RL Chromosoma 104:479-488(1996).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Berkelley;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celisner S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Andrews-Pfannkoch C., Baldwin D.,
 RA Ballieu R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Bernan B.P., Bhandari D., Bolintsov S.,
 RA Borkova D., Botchan M.R., Bouck J., Brockstein P., Bottlier P.,
 RA Burris K.C., Busan D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.W., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos K., Dopp L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hoston D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Jaislin M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Laspo P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshirefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusseum D.R., Pacle J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,

RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spiedling A.C., Stapleton M., Strong R., Sun E.,
RA Switka R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Massaman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodman T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan W., Zhang G., Zhao Q., Zheng L.,
RA Zhang X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,
RT "The genome sequence of *Drosophila melanogaster*.";
RL Science 287:2185-2195 (2000).
CC -1- FUNCTION: BINDS TO SINGLE AND DOUBLE STRANDED DNA AND EXHIBITS
CC DNA-DEPENDENT ATPASE ACTIVITY. UNDERWINDS DUPLEX DNA (BY
CC SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Nucleus (Probable).
CC -1- SIMILARITY: BELONGS TO THE RECA FAMILY. RAD51 SUBFAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: D37788; BAA07039.1; -
CC EMBL: D17726; BAA04580.1; -
CC EMBL: L41342; AAG4873.1; -
CC EMBL: AE003772; AAF57005.1; -
CC HSSP: Q06609; 1B22.
CC FLYBASE: FBgn0011700; Rad51.
CC InterPro: IPR003593; AAA_ATPase.
CC InterPro: IPR000445; HHH.
CC InterPro: IPR003583; HHH 1.
CC InterPro: IPR001553; RecA.
CC Pfam: PF00663; HHH; 1.
CC SMART: SM00382; AAA; 1.
CC SMART: SM00278; HHH1; 1.
CC PROSITE: PS50162; RECA_2; 1.
CC PROSITE: PS50163; RECA_3; 1.
CC DNA-binding: ATP-binding; Nuclear protein.
CC NP_BIND 124 131 ATP (POTENTIAL).
CC SEQUENCE 336 AA; 36647 MW; F99B2B1405B15DB0 CRC64;
SQ
Query Match 39.3%; Score 42; DB 1; Length 336;
Best Local Similarity 42.1%; Pred. No. 15;
Matches 8; Conservative 5; Mismatches 6; Indels 0; Gaps 0;
OY 3 NYIFRAHDSVQTVIAKI 21
Db 185 NVAFTRAHNSDQQTILIKM 203

RESULT 13
NOS_LYMST STANDARD; PRT; 1153 AA.
ID NOS_LYMST
AC 061309;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 15-FEB-2003 (Rel. 42, Last annotation update)
DE Nitric-oxide synthase (EC 1.14.13.39) (NOS, type I) (Neuronal NOS)
DE (N-NOS) (nnos).
GN NOS.
OS *Lymanaea stagnalis* (great pond snail).
OC Eukaryota; Metazoa; Mollusca; Gastropoda; Pulmonata; Basommatophora;
OC Lymnaeidae; Lymnaeidae; Lymnaea.
OX NCBI_TaxID=6523;
OX [1]
RN SEQUENCE FROM N.A. (ISOFORMS LONG AND SHORT).
RC TISSUE=CNS;
RX MEDLINE=98211896; Pubmed=9552167;
RA Korneev S.A., Piper M.R., Picot J., Phillips R., Korneeva E.I.,
RA O'Shea M.,
RT "Molecular characterization of NOS in a mollusc: expression in a giant

RT modulatory neuron.";
RL J. Neurobiol. 35:65-76 (1998).
CC -1- FUNCTION: PRODUCES NITRIC OXIDE (NO) WHICH IS A MESSENGER MOLECULE
CC WITH DIVERSE FUNCTIONS THROUGHOUT THE BODY (BY SIMILARITY).
CC -1- CATALYTIC ACTIVITY: L-arginine + NADPH + H⁺ (2) = citrulline +
CC nitric oxide + NADP(+).
CC -1- COFACTOR: HEME. BINDS ONE MOLE EACH OF FAD AND FMN (BY
CC SIMILARITY).
CC -1- ENZYME REGULATION: STIMULATED BY CALCIUM/CALMODULIN.
CC -1- ALTERNATIVE PRODUCTS:
CC Event-Alternative splicing; Named isoforms=2;
CC Name=long;
CC IsoId=061309-1; Sequence=Displayed;
CC Name=short;
CC IsoId=061309-2; Sequence=VSP 003584;
CC -1- TISSUE SPECIFICITY: EXPRESSED IN THE CENTRAL NERVOUS SYSTEM, IN
CC THE SEROTONERGIC CEREBRAL GIANT CELLS, BOTH THE LONG AND SHORT
CC ISOFORMS ARE EXPRESSED EQUALLY IN THE CNS.
CC -1- SIMILARITY: BELONGS TO THE NOS FAMILY.
CC -1- SIMILARITY: Contains 1 flavodoxin-like domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AF012531; AAC17487.1; -
CC FPIR: T31080; T31080.
CC HSSP: P29477; 1DB7.
CC InterPro: IPR003097; FAD_binding.
CC InterPro: IPR001094; Flavodoxin like.
CC InterPro: IPR001226; Flavodoxin.
CC InterPro: IPR001709; FPN_Cyt_redctse.
CC InterPro: IPR004030; NO_synthase.
CC InterPro: IPR001433; Oxred_FAD/NAD(P).
CC Pfam: PF00667; FAD_binding_1; 1.
CC Pfam: PF00258; Flavodoxin; 1.
CC Pfam: PF00175; NAD_binding_1; 2.
CC Pfam: PF02898; NO_synthase; 1.
CC PRINTS: PR00369; FLAVODOXIN.
CC PRINTS: PR00371; FPNCR.
CC DR PROSITE: PS50902; FLAVODOXIN LIKE; 1.
CC DR PROSITE: PS60001; NOS; 1.
CC KM Oxidoreductase; NADP; FAD; FMN; Calmodulin-binding; Heme; Repeat;
CC KW Alternative splicing;
CC FT DOMAIN 427 610
CC FT METAL 82 82
CC FT DOMAIN 397 417
CC FT NP_BIND 556 587
CC FT NP_BIND 697 708
CC FT NP_BIND 836 846
CC FT NP_BIND 911 929
CC FT NP_BIND 1089 1104
CC FT DOMAIN 934 1010
CC FT VARSPLIC 276 309
CC FT FT
CC FT SEQUENCE 1153 AA; 129085 MW; 101B77D02B6B109 CRC64;
SQ
Query Match 39.3%; Score 42; DB 1; Length 1153;
Best Local Similarity 38.1%; Pred. No. 57;
Matches 8; Conservative 7; Mismatches 6; Indels 0; Gaps 0;
OY 1 MANYIFRAHDSVQTVIAKI 21
Db 667 LAERIQAKDSQQTILIKL 687

RESULT 14
KPVA_RICCO

ID KEPA, RICCO STANDARD; PRT; 583 AA.
 AC Q43117; Q43118;
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Pyruvate Kinase Isozyme A, chloroplast precursor (EC 2.7.1.40).
 OS Ricinus communis (Castor bean).
 OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC Eucosids I; Malviales; Euphorbiaceae; Ricinus.
 NCBI_TaxID=3988;
 [1]
 SEQUENCE FROM N.A. (ISOFORMS ALPHA AND BETA), AND SEQUENCE OF 75-81.
 RC STRAIN=cv. Baker 296; TISSUE=Endosperm;
 RA Blakeley S.D., Plaxton W.C., Dennis D.T.;
 RT Relationship between the subunits of leucoplast pyruvate kinase from
 RT Ricinus communis and a comparison with the enzyme from other
 RT sources.";
 RL Plant Physiol. 96:1283-1288 (1991).
 CC -1- CATALYTIC ACTIVITY: ATP + pyruvate = ADP + phosphoenolpyruvate.
 CC -1- COFACTOR: REQUIRES MAGNESIUM AND POTASSIUM (BY SIMILARITY).
 CC -1- PATHWAY: GLYCOLYSIS; final step.
 CC -1- SUBUNIT: OLIGOMER OF ALPHA AND BETA SUBUNITS (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: Chloroplast.
 CC -1- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=2;
 CC Name=Alpha;
 CC IsoId=Q43117-1; Sequence=Displayed;
 CC Name=Beta;
 CC IsoId=Q43117-2; Sequence=VSP 002886;
 CC -1- SIMILARITY: BELONGS TO THE PYRUVATE KINASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC DR EMBL, M64736; AAA33870.1; -
 CC EMBL, M64737; AAA33871.1; -
 CC PIR, T10051; T10051.
 CC PIR, T10054; T10054.
 CC HSSP, P14178; 1E0T.
 CC DR InterPro: IPR001697; Pyruvate_kinase.
 CC Pfam, PR00224; PK.1.
 CC Pfam, PR02887; PK.C.1.
 CC DR PRINTS: PR01050; PYRUVATE_KINASE.
 CC DR ProDom: PD001009; Pyruvate_kinase; 1.
 CC TRIGRAM: TIGR01064; Pyruv_kin.1.
 CC DR PROSITE: PS00110; PYRUVATE_KINASE; 1.
 CC KW Pyruvate; Transferase; Kinase; Glycolysis; Magnesium; Chloroplast;
 KW Transic peptide; Alternative splicing
 FT TRANSIT 1 74 CHLOROPLAST.
 FT CHAIN 75 583 PYRUVATE KINASE ISOZYME A.
 FT ACT SITE 331 331 BY SIMILARITY.
 FT METAL 333 333 MAGNESIUM (POTENTIAL).
 FT METAL 354 354 MAGNESIUM (POTENTIAL).
 FT METAL 355 355 MAGNESIUM (POTENTIAL).
 FT VARSPIC 1 129 MQQSIFSPNLTFAKQPFKPLPPTPSRYPVANNYSLS
 FT IKASTISFSSSSSPQVVAANGNGNSGVLYNNNKAVTSDP
 FT SSIADVAVTELEKENGFRSTRKTKVCTIGRTGCEFLD
 FT ALAVGG -> MAVVVDLEAVRVVAVLALRMEVVAVLT
 FT AVNGVGD -> (in isoform Beta).
 FT FTID=VSP 002886.
 FT SEQUENCE 583 AA; 64093 MW; 5DFDE4F78A755C2 CRC64;
 Query Match 38.8%; Score 41.5; DB 1; Length 583;
 Best Local Similarity 47.6%; Pred. No. 33;
 Matches 10; Conservative 5; Mismatches 5; Indels 1; Gaps 1;

Db : : : : :
 313 LKSYIAKRSDDI-AVIKI 332
 RESULT 15
 RH23_SCHPO
 ID RH23_SCHPO STANDARD; PRT; 368 AA.
 AC 074803;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE UV excision repair protein rhp23 (RAD23 homolog).
 GN RHP23 OR SPAC2D10.12.
 OS Schizosaccharomyces pombe (fission yeast).
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;
 OC Schizosaccharomycetes.
 NCBI_TaxID=4896;
 [1]
 SEQUENCE FROM N.A., AND SUBCELLULAR LOCATION.
 RP STRAIN=SP223;
 RC MEDLINE=21648918; PubMed=11788722;
 RX Elder R.T., Song X.-Q., Chen M., Hopkins K.M., Lieberman H.B.,
 RA Zhao Y.;
 RA "Involvement of rhp23, a Schizosaccharomyces pombe homolog of the
 RT human HHR23A and Saccharomyces cerevisiae Rad23 nucleotide excision
 RT repair genes, in cell cycle control and protein ubiquitination.";
 RL Nucleic Acids Res. 30:581-591 (2002).
 [2]
 SEQUENCE FROM N.A.
 RP STRAIN=972;
 RC MEDLINE=21848401; PubMed=11859360;
 RX Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
 RA Sgouros J., Peat N., Hayes C., Baker S., Basham D., Bowman S.,
 RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
 RA Collins M., Connor R., Cronin A., Davis P., Feltham T., Fraser A.,
 RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
 RA Holroyd S., Hornsby T., Howarth S., Huckle E.D., Hunt S., Jagers K.,
 RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
 RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
 RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitch E.,
 RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
 RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
 RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
 RA Woodward J., Voickert G., Aert R., Roben J., Grymoprez B.,
 RA Weljens I., Vanstreels E., Rieger M., Schefer M., Mueller-Auer S.,
 RA Gabel C., Fuchs M., Filz C., Holzer E., Moestl D., Hilbert H.,
 RA Borzym K., Langer I., Beck A., Lehrach H., Reinhardt R., Pohl T.M.,
 RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
 RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
 RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
 RA Lucas M., Rochet M., Galliard C., Tallada V.A., Garzon A., Thode G.,
 RA Daga R.R., Cruzado L., Jimenez U., Sanchez J., del Rey F., Benito J.,
 RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsberg S.L.,
 RA Cerrutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
 RA Shpakovski G.V., Ussery D., Bartell B.G., Nurse P.;
 RA "The genome sequence of Schizosaccharomyces pombe.";
 RL Nature 415:871-880 (2002).
 [3]
 CHARACTERIZATION.
 RP MEDLINE=20119127; PubMed=10652237;
 RA Lombaerts M., Goeloe J.I., den Dulk H., Brandsma J.A., Brouwer J.;
 RT "Identification and characterization of the rhp23(+) DNA repair gene
 RT in Schizosaccharomyces pombe.";
 RL Biochem. Biophys. Res. Commun. 268:210-215 (2000).
 CC -1- FUNCTION: Involved in postreplication repair of UV-damaged DNA.
 CC strand on replication repair functions in gap-filling of a daughter
 CC -1- SUBCELLULAR LOCATION: Nuclear.
 CC -1- SIMILARITY: Contains 1 ubiquitin-like domain.
 CC -1- SIMILARITY: Contains 2 UBA domains.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC -----
DR EMBL; AF174293; AAD51975.1; -
DR EMBL; AL031788; CAA21170.1; -
DR PIR; T40115; T40115.
DR HSSP; P54725; IDV0.
DR GeneDB Spombe; SPBC2D10.12; -
DR InterPro; IPR004806; Rad23.
DR InterPro; IPR006636; STI1.
DR InterPro; IPR000449; UBA_domain.
DR InterPro; IPR000626; Ubiquitin.
DR Pfam; PF00627; UBA; 2.
DR Pfam; PF00240; ubiquitin; 1.
DR SMART; SM00727; STI1; 1.
DR SMART; SM00165; UBA; 2.
DR SMART; SM00213; UBQ; 1.
DR TIGRFAMs; TIGR00601; rad23; 1.
DR PROSITE; PSS0030; UBA; 2.
DR PROSITE; PSS0053; UBQUITIN 2; 1.
DR DNA damage; DNA repair; Nuclear protein; Repeat.
KM DOMAIN 1 77 UBQUITIN-LIKE.
FT DOMAIN 135 185 UBA 1.
FT DOMAIN 320 360 UBA 2.
FT DOMAIN 119 122 POLY-ALA.
FT DOMAIN 205 208 POLY-GLN.
SQ SEQUENCE 368 AA; 40135 MW; SCE75EB7E190EFPD CRC64;

Query Match 38.3%; Score 41; DB 1; Length 368;
Best Local Similarity 58.3%; Pred. No. 25;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 ANYFIRAHDE 13
Db 354 ANYLFEHQHSE 365

Search completed: November 13, 2003, 09:45:34
Job time : 8.92417 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 13, 2003, 09:31:40 ; Search time 33.1564 Seconds
(without alignments)
171.224 Million cell updates/sec

Title: US-09-290-049a-16
Sequence: 107
1 MANYIFIRAHDSFVQTIAKII 22

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 segs, 258052604 residues
Total number of hits satisfying chosen parameters: 830525

Minimum DB seg length: 0
Maximum DB seg length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	100	93.5	1575	2	091CH3
2	95	88.8	1577	2	054178
3	90	84.1	1338	2	09WXJ4
4	89	83.2	1577	2	055265
5	89	83.2	2835	2	08G9Q2
6	88	82.2	1449	2	068542
7	88	82.2	1449	2	055264
8	87	81.3	1016	2	091CJ7
9	87	81.3	1477	2	091466
10	87	81.3	1508	2	09EZHS
11	87	81.3	1508	2	052224
12	87	81.3	1512	2	09WXJ5
13	87	81.3	1599	2	000599
14	86	80.4	1527	2	09ZAR4
15	86	80.4	1527	2	08KRE1
16	85	79.4	1290	2	048756

17	84	78.5	1554	2	08KZL5	08KZL5 streptococc
18	78	72.9	1518	2	000600	000600 streptococc
19	64	59.8	1590	2	059383	059383 streptococc
20	64	59.8	1590	2	055263	055263 streptococc
21	53	49.5	2057	2	09RE05	09RE05 leucostoc
22	48.5	45.3	771	16	09EDZ6	09EDZ6 shewanella
23	47	43.9	1197	5	09B121	09B121 drosophila
24	47	43.9	1197	5	09VCH8	09VCH8 drosophila
25	47	43.9	1387	5	0964N5	0964N5 aplysia cal
26	46.5	43.5	783	5	08SX46	08SX46 drosophila
27	46.5	43.5	866	5	09V8T7	09V8T7 drosophila
28	46	43.0	881	16	08G5W2	08G5W2 bifidobacte
29	45	42.1	368	16	08K610	08K610 streptococc
30	45	42.1	437	16	099YF6	099YF6 streptococc
31	45	42.1	437	16	08P020	08P020 streptococc
32	45	42.1	533	10	08H6T2	08H6T2 chlamydomon
33	45	42.1	665	5	09VAL1	09VAL1 drosophila
34	45	42.1	805	16	08DDJ2	08DDJ2 vibrio vuln
35	44	41.1	132	16	09KPH6	09KPH6 vibrio chol
36	44	41.1	264	16	08D618	08D618 vibrio vuln
37	44	41.1	341	2	09L7T7	09L7T7 rhodospirill
38	44	41.1	583	10	049505	049505 aradidopsis
39	44	41.1	694	2	0915M3	0915M3 salmonella
40	44	41.1	694	16	0935G9	0935G9 salmonella
41	44	41.1	1213	3	059801	059801 schistosach
42	43	40.2	261	17	P94945	P94945 methanopyru
43	43	40.2	288	5	09N8M7	09N8M7 caenorhabdi
44	43	40.2	933	16	051486	051486 dorrellia bu
45	42	39.3	214	16	08ED17	08ED17 shewanella

ALIGNMENTS

RESULT 1

Q91CH3 ID Q91CH3 PRELIMINARY; PRT; 1575 AA.

AC Q91CH3; 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-OCT-2002 (TREMBLrel. 22, Last annotation update)
DE Glucosyltransferase.
GN GTRF.
OS Streptococcus oralis.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1303;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC10557; PubMed=10768934;
RX MEDLINE=20231779; Pubmed=10768934;
RA Fujiwara T., Hoshino T., Ooshima T., Sobue S., Hamada S.;
RT "Purification, characterization, and molecular analysis of the gene
encoding glucosyltransferase from Streptococcus oralis.";
RT Infect. Immun. 68:2475-2483(2000).
FI EMBL; AB025228; BAA95201.1; --
DR EMBL; AB025228; BAA95201.1; --
DR InterPro; IPR002479; CW_binding.
DR InterPro; IPR003318; Glyco_hydro_70.
DR Pfam; PF01473; CW_binding_1; 17.
DR Pfam; PF02324; Glyco_hydro_70; 1.
KW Transferase.
SQ SEQUENCE 1575 AA; 176792 MW; 772A2654D7C2B543 CRC64;

Query Match 93.5%; Score 100; DB 2; Length 1575;
Best Local Similarity 90.9%; Pred. No. 7.2e-08;
Matches 20; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 MANYIFIRAHDSFVQTIAKII 22
|||
DB 617 MANYIFIRAHDSFVQTIAKII 638

RESULT 2

```

Q54178      PRELIMINARY;      PRT; 1577 AA.
ID   Q54178; Q54247;
AC   Q54178; Q54247;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   Glucosyltransferase.
GN   GTFG.
OS   Streptococcus gordonii Challis.
OC   Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC   Streptococcus.
OX   NCBI_TaxID=29390;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CHALLIS;
RA   MEDLINE=96157084; PubMed=8586195;
RA   Vickerman M.M.; Sulaivik M.C.; Clewell D.B.;
RT   "Molecular analysis of Streptococcus gordonii glucosyltransferase
RT   phase variants";
RL   Dev. Biol. Stand. 85:309-314(1995).
RN   [2]
RP   SEQUENCE OF 1-96 FROM N.A.
RC   STRAIN=CHALLIS;
RX   MEDLINE=92276337; PubMed=1534326;
RA   Sulaivik M.C.; Tardif G.; Clewell D.B.;
RT   "Identification of a gene, rgg, which regulates expression of
RT   glucosyltransferase and influences the Spp phenotype of Streptococcus
RT   gordonii Challis.";
RL   J. Bacteriol. 174:3577-3586(1992).
DR   EMBL; U12643; AAC3483.1; -
DR   EMBL; M89776; AAA26969.1; -
DR   InterPro; IPR002479; CW binding.
DR   InterPro; IPR003318; Glyco_hydro_70.
DR   Pfam; PF01473; CW_binding_1; 18.
DR   Pfam; PF02324; Glyco_hydro_70; 1.
KW   Transferase.
SQ   SEQUENCE 1577 AA; 177805 MW; 5AE0328DC5E08D18 CRC64;

Query Match      88.8%; Score 95; DB 2; Length 1577;
Best Local Similarity 90.5%; Pred. No. 5.1e-07;
Matches 19; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY      2 ANYIFRAHDSVQTVIAKII 22
DB      620 ANYIFRAHDSVQTVIAKII 640

RESULT 3
ID   Q9WKJ4      PRELIMINARY;      PRT; 1338 AA.
AC   Q9WKJ4;
DT   01-NOV-1999 (TrEMBLrel. 12, Created)
DT   01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   GTF-S.
GN   GTF-S.
OS   Streptococcus criceti.
OG   Plasmid pAK1.
OC   Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC   Streptococcus.
OX   NCBI_TaxID=1333;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=HS-6;
RA   Inoue M.; Fukui K.; Miyagi A.;
RT   "S.cricetus glucosyltransferase (gfts and gtfI) genes";
RT   Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
DR   EMBL; AB026123; BAA77236.1; -
DR   InterPro; IPR002479; CW binding.
DR   InterPro; IPR003318; Glyco_hydro_70.
DR   Pfam; PF01473; CW_binding_1; 10.
DR   Pfam; PF02324; Glyco_hydro_70; 1.
KW   Plasmid.

```

```

SQ   SEQUENCE 1338 AA; 148558 MW; 0A90C8E10E15D39B CRC64;

Query Match      84.1%; Score 90; DB 2; Length 1338;
Best Local Similarity 90.0%; Pred. No. 3e-06;
Matches 18; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY      3 ANYIFRAHDSVQTVIAKII 22
DB      511 NYIFRAHDSVQTVIAKII 530

RESULT 4
ID   Q55265      PRELIMINARY;      PRT; 1577 AA.
AC   Q55265;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   Glucosyltransferase precursor.
GN   GTFM.
OS   Streptococcus salivarius.
OC   Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC   Streptococcus.
OX   NCBI_TaxID=1304;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   MEDLINE=95122197; PubMed=7822030;
RA   Simpson C.L.; Giffard P.M.; Jacques N.A.;
RT   "Streptococcus salivarius ATCC 25975 possesses at least two genes
RT   coding for primer-independent glucosyltransferases.";
RL   Infect. Immun. 63:609-621(1995).
DR   EMBL; L35928; AAC41413.1; -
DR   InterPro; IPR004828; Cutinase_antigen.
DR   InterPro; IPR002479; CW_binding.
DR   InterPro; IPR003318; Glyco_hydro_70.
DR   Pfam; PF01473; CW_binding_1; 11.
DR   Pfam; PF02324; Glyco_hydro_70; 1.
DR   ProDom; PD15343; Cutinase_antigen; 1.
KW   Signal; Transferase.
FT   CHAIN      1..38      POTENTIAL.
FT   SIGNAL     1          GLUCOSYLTRANSFERASE.
FT   CHAIN      39..1577   3EPB89BA7D3AVBF3 CRC64;
SQ   SEQUENCE 1577 AA; 175290 MW; 3EPB89BA7D3AVBF3 CRC64;

Query Match      83.2%; Score 89; DB 2; Length 1577;
Best Local Similarity 81.0%; Pred. No. 5.4e-06;
Matches 17; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY      2 ANYIFRAHDSVQTVIAKII 22
DB      660 ANYIFRAHDSVQTVIAKII 680

RESULT 5
ID   Q8G9Q2      PRELIMINARY;      PRT; 2835 AA.
AC   Q8G9Q2;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Dextranucrase (EC 2.4.1.5) (Fragment).
GN   DSRE.
OS   Leuconostoc mesenteroides.
OC   Bacteria; Firmicutes; Lactobacillales; Leuconostoc.
OX   NCBI_TaxID=1245;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   MEDLINE=22231661; PubMed=12270834;
RA   Bozonnet S.; Dols-Lafargue M.; Fabre E.; Pizunt S.; Renaud-Simeon M.;
RA   Mondan P.; Willemot R.M.;
RT   "Molecular characterisation of DSR-E, an alpha-1,2 linkage
RT   synthesising dextranucrase with two catalytic domains.";
RL   J. Bacteriol. 184:5753-5761(2002).
DR   EMBL; A430204; CAD22883.1; -

```

```
KM Transferase; Glycosyltransferase.
FT NON TER      1
SQ SEQUENCE     2835 AA; 313264 MW; D03262CD735399D CRC64;

Query Match
Best Local Similarity 83.2%; Score 89; DB 2; Length 2835;
Matches 18; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY          3 NYIFIRAHDSVQTVIAKII 22
              ||| ||||| ||||| |||
Db          630 NYAFIRAHDSVQTIVIAQII 649

RESULT 6
O68542       PRELIMINARY; PRT; 1449 AA.
ID O68542;
AC O68542;
DT 01-AUG-1998 (TrEMBLrel. 07, Created)
DR 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Glucosyltransferase N (Fragment).
GN GFEN.
OS Streptococcus salivarius.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
CX NCBI_TaxID=1304;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=V1477;
RA Jaffe R.I.;
RT "Streptococcus salivarius V1477 GFEN.";
RL Submitted (FEB-1998) to the EMBL/genbank/DDBJ databases.
DR EMBL; AF049609; AAC05156.1; -.
DR InterPro; IPR002479; CW binding.
DR InterPro; IPR003318; Glyco_hydro_70.
DR Pfam; PF01473; CW_binding_I; 8.
DR Pfam; PF02324; Glyco_hydro_70; 1.
KW Transferase.
KM NON TER      1449
FT FT           1449
SQ SEQUENCE     1449 AA; 159895 MW; 0700FD748471BFB CRC64;

Query Match
Best Local Similarity 82.2%; Score 88; DB 2; Length 1449;
Matches 16; Conservative 4; Mismatches 2; Indels 0; Gaps 0

QY          1 MANYIFIRAHDSVQTVIAKII 22
              ||| ||||| ||||| |||
Db          607 MANYAFIRAHDSVQSIIIGIIT 628

RESULT 7
O55264       PRELIMINARY; PRT; 1449 AA.
ID O55264;
AC O55264;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DR 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Glucosyltransferase precursor.
GN GFPL.
OS Streptococcus salivarius.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
CX NCBI_TaxID=1304;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=95122197; PubMed=7822030;
RA Simpson C.L., Giffard P.W., Jacques N.A.;
RT "Streptococcus salivarius ATCC 25975 possesses at least two genes coding for primer-independent glucosyltransferases."
RT Infect. Immun. 63:609-621 (1995).
DR EMBL; U35495; AAC41412.1; -.
DR InterPro; IPR002479; CW_binding.
DR InterPro; IPR003318; Glyco_hydro_70.
```

```

DR Pfam: PF01473; CM binding_1; 8.
DR Pfam: PF02324; Glyco_hydro_70; 1.
KM Signal; Transferase.
FT SIGNAL 1 35 POTENTIAL.
FT CHAIN 36 1449 GLUCOSYLTRANSFERASE.
SQ SEQUENCE 1449 AA; 159984 MW; D622F07306E86A46 CRC64;

Query Match 82.2%; Score 88; DB 2; Length 1449;
Best Local Similarity 72.7%; Pred. NO. 7.2e-06;
Matches 16; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

Qy 1 MANYFIRADSEVQTVIAKII 22
||| - ||||| ||||| :: ||
||| - ||||| ||||| :: ||
Db 607 MANYAFVRAHDSVQSIIGQII 628

RESULT 8
O9LCJ7 PRELIMINARY; PRT; 1016 AA.
AC O9LCJ7;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DE 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
DE Dextranucrase.
GN DSRT.
OS Leuconostoc mesenteroides.
OC Bacteria; Firmicutes; Lactobacillales; Leuconostoc.
OX NCBI_TaxID=1245;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NRRL B-512F;
RX MEDLINE=20169623; PubMed=10705445;
RA Punane K., Mizuno K., Takahara H., Kobayashi M.;
RT "Gene encoding a dextranucrase-like protein in Leuconostoc
RL mesenteroides NRRL B-512F.";
RL Biosci. Biotechnol. Biochem. 64:29-38(2000).
DR EMRL; AB020020; BAA90527.1; -.
HSEB; P06278; 1V4S.
DR InterPro; IPR00318; Glyco_hydro_70.
DR Pfam; PF02324; Glyco_hydro_70; 1.
SQ SEQUENCE 1016 AA; 110344 MW; 8896EFD813CCB47 CRC64;

Query Match 81.3%; Score 87; DB 2; Length 1016;
Best Local Similarity 85.0%; Pred. NO. 7.3e-06;
Matches 17; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 3 NYIFIRADSEVQTVIAKII 22
||| : ||||| ||||| :: ||
||| : ||||| ||||| :: ||
Db 626 NYSFVRAHDSVQTVIAEII 645

RESULT 9
O91466 PRELIMINARY; PRT; 1477 AA.
AC O91466;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DE 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
DE Dextranucrase (EC 2.4.1.5).
GN DSRC.
OS Leuconostoc mesenteroides.
OC Bacteria; Firmicutes; Lactobacillales; Leuconostoc.
OX NCBI_TaxID=1245;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NRRL B-1355;
RA Arguello-Morales M.A., Remaud-Simeon M., Pizut S., Sarcabal P.,
RA Willemot R.M., Monsan P.;
RT "Sequence analysis of the gene encoding alternanucrase, a sucrose
RT glucosyltransferase from Leuconostoc mesenteroides NRRL B-1355.";
RL Submitted (OCT-1999) to the EMBL/GenBank/DBD databases.
EMBL; AJ250172; CAB76565.1; -.
DR InterPro; IPR002479; CM_binding.

```

DR InterPro: IPR003318; Glyco_hydro_70.
 DR Pfam: PF01473; CW_binding_1; 14.
 DR Pfam: PF02324; Glyco_hydro_70; 1.
 KW Glycosyltransferase; Transferase.
 SQ SEQUENCE 1477 AA; 164887 MW; E6F5710DEFCB831 CRC64;

Query Match 81.3%; Score 87; DB 2; Length 1477;
 Best Local Similarity 85.0%; Pred. No. 1.1e-05;
 Matches 17; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 3 NYIFIRAHDSVQTVIAKII 22
 Db 605 NYSFVRAHDSVQTVIAQII 624

RESULT 10

Q9EZH5 PRELIMINARY; PRT; 1508 AA.
 AC Q9EZH5;
 DT 01-MAR-2001 (TrEMBLrel. 16, Created)
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
 DE Dextranucrase DsrB742.
 GN DSRB742.
 OS Leuconostoc mesenteroides.
 OC Bacteria; Firmicutes; Lactobacillales; Leuconostoc.
 OX NCBI_TaxID=1245;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=B-742CB;
 RA Kim H.-S., Kim D., Ryu H.-J., Robyt J.F.;
 RL "Leuconostoc mesenteroides B-742CB, a dextranucrase gene."
 Submitted (AUG-2000) to the EMBL/Genbank/DBJ databases.
 DR EMBL; AF294469; AAG38021.1;
 DR InterPro: IPR002479; CW_binding.
 DR InterPro: IPR003318; Glyco_hydro_70.
 DR Pfam: PF01473; CW_binding_1; 14.
 DR Pfam: PF02324; Glyco_hydro_70; 1.
 DR Pfam: PF02324; Glyco_hydro_70; 1.
 SQ SEQUENCE 1508 AA; 168542 MW; E2FCFAP087AEF3A CRC64;

Query Match 81.3%; Score 87; DB 2; Length 1508;
 Best Local Similarity 85.0%; Pred. No. 1.1e-05;
 Matches 17; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 3 NYIFIRAHDSVQTVIAKII 22
 Db 636 NYSFVRAHDSVQTVIAQII 655

RESULT 11

OS2224 PRELIMINARY; PRT; 1508 AA.
 AC OS2224;
 DT 01-JUN-1998 (TrEMBLrel. 06, Created)
 DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
 DE Glucosyltransferase (EC 2.4.1.5).
 GN DSRB.
 OS Leuconostoc mesenteroides.
 OC Bacteria; Firmicutes; Lactobacillales; Leuconostoc.
 OX NCBI_TaxID=1245;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=NRRL B-1299;
 RA Monchois V., Renaud-Simeon M., Monnan P., Willemot R.M.;
 RL "Cloning and sequencing of a gene coding for an extracellular
 dextranucrase (DSRB) from Leuconostoc mesenteroides NRRL B-1299
 RT synthesizing only a a(1-6) glucan."
 RT EMBL; AF030129; AAB95453.1;
 DR EMBL; AF030129; AAB95453.1;
 DR InterPro: IPR002479; CW_binding.
 DR InterPro: IPR003318; Glyco_hydro_70.
 DR Pfam: PF01473; CW_binding_1; 14.

DR Pfam: PF02324; Glyco_hydro_70; 1.
 KW Glycosyltransferase; Transferase.
 SQ SEQUENCE 1508 AA; 168511 MW; E70CCEB57A70D1F0 CRC64;

Query Match 81.3%; Score 87; DB 2; Length 1508;
 Best Local Similarity 85.0%; Pred. No. 1.1e-05;
 Matches 17; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 3 NYIFIRAHDSVQTVIAKII 22
 Db 636 NYSFVRAHDSVQTVIAQII 655

RESULT 12

Q9WKJ5 PRELIMINARY; PRT; 1512 AA.
 AC Q9WKJ5;
 DT 01-NOV-1999 (TrEMBLrel. 12, Created)
 DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
 DE GTF.
 GN GTF.
 OS Streptococcus criceti.
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 OX NCBI_TaxID=1333;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=HS-6;
 RA Inoue M., Fukui K., Miyagi A.;
 RL "S.cricetus glucosyltransferase(gtfS and gtfT) genes."
 Submitted (MAR-1999) to the EMBL/Genbank/DBJ databases.
 DR EMBL; AB026123; BAA77237.1;
 DR InterPro: IPR002479; CW_binding.
 DR InterPro: IPR003318; Glyco_hydro_70.
 DR Pfam: PF01473; CW_binding_1; 14.
 DR Pfam: PF02324; Glyco_hydro_70; 1.
 DR Pfam: PF02324; Glyco_hydro_70; 1.
 KW Pfam: PF02324; Glyco_hydro_70; 1.
 SQ SEQUENCE 1512 AA; 167145 MW; 4C03D9C8C601FC14 CRC64;

Query Match 81.3%; Score 87; DB 2; Length 1512;
 Best Local Similarity 80.0%; Pred. No. 1.1e-05;
 Matches 16; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Qy 3 NYIFIRAHDSVQTVIAKII 22
 Db 561 SYFVRAHDSVQTVIAQII 580

RESULT 13

Q00599 PRELIMINARY; PRT; 1599 AA.
 AC Q00599;
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
 DE Glucosyltransferase S precursor (EC 2.4.1.5) (GTF) (Dextranucrase)
 GN (Sucrose 6-glucosyltransferase).
 GN GTFK.
 OS Streptococcus salivarius.
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 OX NCBI_TaxID=1304;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 25975;
 RX MEDLINE=93381463; PubMed=8371114;
 RA Gifford P.M., Allen D.M., Milward C.P., Simpson C.L., Jacques N.A.;
 RL "Sequence of the gtfK gene of Streptococcus salivarius ATCC 25975 and
 RT evolution of the gtf genes of oral streptococci."
 RT J. Gen. Microbiol. 139:1511-1522 (1993).
 RN [2]

RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 25975;
 RX MEDLINE=92148377; PubMed=1838391;
 RA Giffard P.M., Simpson C.L., Milward C.P., Jacques N.A.;
 RT "Molecular characterization of a cluster of at least two
 RT glycosyltransferase genes in *Streptococcus salivarius* ATCC 25975.";
 RL J. Gen. Microbiol. 137:2577-2593 (1991).
 CC -1- FUNCTION: PRODUCTION OF EXTRACELLULAR GLUCANS, THAT ARE THOUGHT TO
 CC PLAY A KEY ROLE IN THE DEVELOPMENT OF THE DENTAL PLAQUE BECAUSE OF
 CC THEIR ABILITY TO ADHERE TO SMOOTH SURFACES AND MEDIATE THE
 CC AGGREGATION OF BACTERIAL CELLS AND FOOD DEBRIS.
 CC -1- CATALYTIC ACTIVITY: SUCROSE + ((1,6)-ALPHA-D-GLUCOSYL) (N) = D-
 CC FRUCTOSE + ((1,6)-ALPHA-D-GLUCOSYL) (N+1).
 CC -1- SUBCELLULAR LOCATION: EXTRACELLULAR.
 CC -1- DISEASE: DENTAL CARIES.
 CC -1- SIMILARITY: TO REGIONS OF BARLEY AND BACILLUS AMYLOLIQUEFACIENS
 CC ALPHA AMYLASES AND RABBIT GLYCOGEN PHOSPHORYLASE.
 DR EMBL; Z11872; CA477898.1; -;
 DR EMBL; Z11873; CA477901.1; -;
 DR EMBL; M6411; AAA26897.1; -;
 DR InterPro; IPR002479; CW_binding.
 DR InterPro; IPR003318; Glyco_hydro_70.
 DR Pfam; PF01473; CW_binding_1; 14.
 DR Pfam; PF02324; Glyco_hydro_70; 1.
 DR Transferrase; Glycosyltransferase; Signal; Repeat; Dental caries.
 FT SIGNAL 1 42 POTENTIAL.
 FT CHAIN 43 1599 GLYCOSYLTRANSFERASE S.
 SQ SEQUENCE 1599 AA; 176480 MW; 24877869E152B707 CRC64;

Query Match 81.3%; Score 87; DB 2; Length 1599;
 Best Local Similarity 81.0%; Pred. No. 1.2e-05;
 Matches 17; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 2 NYIFIRAHSEVQTVIAKII 22
 DB 573 ATYIFVRAHSEVQTVIADII 593

RESULT 14
 O9ZAR4 PRELIMINARY; PRT; 1527 AA.
 AC O9ZAR4;
 DT 01-MAY-1999 (TRENBLREL. 10, Created)
 DT 01-MAY-1999 (TRENBLREL. 10, Last sequence update)
 DT 01-OCT-2002 (TRENBLREL. 22, Last annotation update)
 DE Dextranucrase.
 GN DEX.
 OS Leuconostoc mesenteroides.
 OC Bacteria; Firmicutes; Lactobacillales; Leuconostoc.
 OX NCBI_TaxID=1245;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=NRRL B-512-F;
 RT "Cloning and Molecular Characterization of Dextranucrase Gene from
 RT Leuconostoc mesenteroides NRRL B-512F.";
 RT Submitted (DEC-1996) to the EMBL/GenBank/DBJ databases.
 DR EMBL; U81374; AAD10952.1; -;
 DR InterPro; IPR002479; CW_binding.
 DR InterPro; IPR003318; Glyco_hydro_70.
 DR Pfam; PF01473; CW_binding_1; 16.
 DR Pfam; PF02324; Glyco_hydro_70; 1.
 DR SEQUENCE 1527 AA; 169709 MW; 1DFAFA237C743398 CRC64;

Query Match 80.4%; Score 86; DB 2; Length 1527;
 Best Local Similarity 80.0%; Pred. No. 1.7e-05;
 Matches 16; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 3 NYIFIRAHSEVQTVIAKII 22
 DB 654 NYSFVRAHSEVQTVIAQIV 673

RESULT 15
 O9KRE1 PRELIMINARY; PRT; 1527 AA.
 AC O9KRE1;
 DT 01-OCT-2002 (TRENBLREL. 22, Created)
 DT 01-OCT-2002 (TRENBLREL. 22, Last sequence update)
 DT 01-MAR-2003 (TRENBLREL. 23, Last annotation update)
 DE Dextranucrase DsrD (EC 2.4.1.5).
 GN DSRD.
 OS Leuconostoc mesenteroides.
 OC Bacteria; Firmicutes; Lactobacillales; Leuconostoc.
 OX NCBI_TaxID=1245;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Neubauer H., Bauche A., Mollet B.;
 RT "Isolation and characterization of the dextranucrase DsrD of
 RT Leuconostoc mesenteroides Lcc4.";
 RT Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY017384; AAG61158.1; -;
 DR InterPro; IPR002479; CW_binding.
 DR InterPro; IPR003318; Glyco_hydro_70.
 DR Pfam; PF01473; CW_binding_1; 6.
 DR Pfam; PF02324; Glyco_hydro_70; 1.
 DR Transferrase; Glycosyltransferase.
 SQ SEQUENCE 1527 AA; 169835 MW; F9D0DE220BD89668 CRC64;

Query Match 80.4%; Score 86; DB 2; Length 1527;
 Best Local Similarity 80.0%; Pred. No. 1.7e-05;
 Matches 16; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 3 NYIFIRAHSEVQTVIAKII 22
 DB 654 NYSFVRAHSEVQTVIAQIV 673

Search completed: November 13, 2003, 09:44:09
 Job time : 35.1564 secs